



Université de Sherbrooke

**Caractérisation des fonctions moléculaires des isoformes de NudCD1 dans le cancer**

Par  
Patrick Asselin-Mullen  
Programme de biologie cellulaire

Mémoire présenté à la Faculté de médecine et des sciences de la santé  
en vue de l'obtention du grade de maître ès sciences (M. Sc.)  
en biologie cellulaire

Sherbrooke, Québec, Canada  
Avril 2017

Membres du jury d'évaluation  
Pr François-Michel Boisvert, programme de biologie cellulaire  
Pr Fernand-Pierre Gendron, programme de biologie cellulaire  
Pr Alexandre Maréchal, programme de biologie

©Patrick Asselin-Mullen, 2017

## Caractérisation des fonctions moléculaires des isoformes de NudCD1 dans le cancer

par

Patrick Asselin-Mullen

Programme de biologie cellulaire

Mémoire présenté à la Faculté de médecine et des sciences de la santé en vue de l'obtention du diplôme de maître ès sciences (M.Sc.) en biologie cellulaire, Faculté de Médecine et des Sciences de la Santé, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1H 5N4.

Le cancer est la cause principale de décès chez les canadiens. Cette maladie est souvent caractérisée par des surexpressions, des sous-expressions et des mutations chez certaines protéines clés. NudCD1 est un gène ayant été identifié lors de la dernière décennie comme possédant une forte expression chez des patients atteints de leucémie myéloïde chronique. La protéine issue de ce gène n'est normalement exprimée qu'au niveau des tissus sains testiculaires et cardiaques, elle est cependant surexprimée dans différents cancers. La forte expression de NudCD1 est associée avec des phénotypes tumoraux, possiblement agissant sur l'activation d'IGF1R qui permettrait les phosphorylations subséquentes des voies de signalisation passant par AKT et ERK1/2. Le gène de NudCD1 est constitué de 12 exons pouvant être épissés alternativement de manière à produire 3 différentes isoformes. Ces isoformes possèdent une région C-terminale commune de 492 acides aminés et ont une région N-terminale spécifique. Les isoformes 2 et 3 ne sont présentement que peu caractérisées et leur rôle potentiel ou présence dans le cancer est inconnu. Cette étude nous suggère que **les différentes isoformes de NudCD1 possèdent des fonctions différentes dans le cancer**. L'expression de ces différentes isoformes a été mesurée au niveau des transcrits d'ARNm ainsi qu'au niveau des protéines dans différentes lignées cellulaires cancéreuses et non-cancéreuses, principalement de types colorectales. NudCD1-1 est exprimée fortement au niveau transcriptionnel et protéique, l'isoforme 2 n'est pas exprimée dans les lignées cellulaires étudiées alors que la troisième isoforme est présente faiblement au niveau d'ARNm chez la majorité des lignées étudiées. La stabilité protéique des isoformes a également été mesurée et démontre que NudCD1-2 et NudCD1-3 sont dégradées rapidement par le protéasome. L'observation microscopique des isoformes par immunofluorescence a démontré que les isoformes n'ont pas la même localisation cellulaire, la première étant principalement nucléaire, la deuxième majoritairement cytoplasmique et la troisième est pancellulaire. Des expériences de spectrométrie de masse en tandem couplées à la méthode de quantification SILAC ont démontré l'interaction entre la protéine NudCD1-1 et DHX15. Ces mêmes expériences ont montré que NudCD1-2 pourrait interagir au niveau du cytosol avec HSP90 dans la réponse aux chocs thermiques et le mouvement nucléaire alors que NudCD1-3 pourrait également agir avec ces protéines en plus de jouer un rôle dans la transcription. L'interaction NudCD1-1 et DHX15 suggère que la protéine pourrait agir au niveau du spliceosome ainsi que la transcription afin de médier l'activation d'IGF1R pour mener au phénotype tumoral. Ces résultats suggèrent que les différentes isoformes ont des fonctions différentes dans la cellule, mais ne sont pas toutes impliquées dans le cancer, en plus d'approfondir les connaissances sur l'action de NudCD1-1 dans le cancer.

Mots clés : Spectrométrie de Masse, NudCD1, DHX15, Cancer

## TABLE DES MATIÈRES

<b>1. Introduction.....</b>	<b>1</b>
1.1 Transcription et épissage.....	1
1.1.1 Synthèse du pré-ARNm.....	1
1.1.2 Épissage du pré-ARNm.....	3
1.1.3 Épissage alternatif du pré-ARNm.....	6
1.1.4 Régulation de l'épissage.....	7
1.1.5 Relations transcription et épissage.....	8
1.2. Cancer.....	11
1.2.1 Voie de signalisation IGF1R.....	13
1.2.2 Voie de signalisation AKT.....	15
1.2.3 Voie de signalisation des MAPK.....	16
1.3. NudCD1.....	17
1.3.1 Famille Nud.....	17
1.3.2 NudC.....	19
1.3.3 NudCD1.....	21
1.3.4 Isoformes de NudCD1.....	22
1.4 Problématique de recherche.....	23
<b>2. Matériel et méthode.....</b>	<b>25</b>
2.1 Culture cellulaire.....	25
2.2 Lignées cellulaire stables, clonages et RT-PCR.....	26
2.3 Transfection transitoire et stable.....	27
2.4 SDS-PAGE et immunobuvardage.....	28
2.5 Immunofluorescence.....	29
2.6 Culture cellulaire par SILAC.....	30
2.7 Immunoprécipitations et précipitations par affinité.....	30
2.8 Digestion protéique.....	31
2.8.1 Digestion sur gel des protéines étiquetées GFP.....	31
2.8.2 Digestion sur billes des protéines étiquetées BirA.....	32
2.9 Chromatographie liquide couplée à la spectrométrie de masse en tandem (LC-MS/MS).....	32

2.10 Quantification et analyse bio-informatique.....	33
<b>3. Résultats.....</b>	<b>35</b>
3.1. Niveaux d'expression des trois isoformes de NudCD1 dans différentes lignées cellulaires .....	35
3.1.1 Par méthode de RT-PCR.....	35
3.1.2 Par méthode d'immunobuvardage .....	36
3.2. Génération de lignées stables de chaque isoforme de NudCD1.....	37
3.2.1 Lignées stables avec étiquette GFP .....	37
3.2.2 Lignées stables avec étiquette BirA.....	38
3.3. Dégradation des isoformes de NudCD1 par le protéasome .....	39
3.4. Localisation des isoformes de NudCD1 dans la cellule .....	40
3.4.1 Étiquette GFP N-terminale .....	40
3.4.2 Étiquette GFP C-terminale.....	41
3.5. Partenaires d'interaction de NudCD1 .....	42
3.6. La protéine DHX15 interagit avec l'isoforme 1 de NudCD1 .....	47
3.7. Partenaires d'interaction de DHX15 .....	48
3.8. NudCD1 colocalise avec DHX15 .....	51
<b>4. Discussion .....</b>	<b>52</b>
4.1 Les isoformes de NudCD1 ne sont pas toutes exprimées dans les différentes lignées cellulaires cancéreuses. ....	52
4.2 Les isoformes 2 et 3 sont dégradées rapidement par le protéasome par rapport à l'isoforme 1 .....	54
4.3 Les 3 isoformes de NudCD1 sont localisées à différents endroits de la cellule....	55
4.4 Partenaires d'interaction des différentes isoformes de NudCD1 .....	56
4.5 L'isoforme 1 de NudCD1 interagit avec la protéine DHX15 .....	63
<b>5. Conclusion .....</b>	<b>68</b>
<b>6. Liste des références.....</b>	<b>71</b>
<b>7. Annexe .....</b>	<b>83</b>
7.1 Protein interaction network of alternatively spliced NudCD1 isoforms .....	83
7.2 Amorces utilisées pour les RT-PCR.....	104

## LISTE DES TABLEAUX

<b>Tableau 1:</b>	Lignées cellulaires utilisées.....	<b>26</b>
<b>Tableau 2:</b>	Vecteurs d'expression utilisés.....	<b>27</b>
<b>Tableau 3:</b>	Anticorps primaire et secondaire utilisés .....	<b>29</b>
<b>Tableau 4 :</b>	Liste des protéines interagissant avec chacune des isoformes de NudCD1-GFP analysées au spectromètre de masse avec un ratio supérieur à 2,0 .....	<b>45</b>
<b>Tableau 5:</b>	Liste des protéines interagissant avec chacune des isoformes de NudCD1-BirA analysées au spectromètre de masse avec un ratio supérieur à 2,0 .....	<b>47</b>
<b>Tableau 6:</b>	Liste des protéines interagissant avec DHX15 identifiées par spectrométrie de masse .....	<b>50</b>

## LISTE DES FIGURES

<b>Figure 1:</b>	Synthèse et modifications du pré-ARNm.....	<b>3</b>
<b>Figure 2:</b>	Épissage du pré-ARNm par le spliceosome .....	<b>5</b>
<b>Figure 3:</b>	Transestérification du complexe C .....	<b>6</b>
<b>Figure 4:</b>	Types d'épissage alternatif.....	<b>7</b>
<b>Figure 5:</b>	Modèles de répression de l'épissage par hnRNP A1 .....	<b>9</b>
<b>Figure 6:</b>	Voie de signalisation IGF1R .....	<b>15</b>
<b>Figure 7:</b>	Les protéines Nud et leurs interactions .....	<b>19</b>
<b>Figure 8:</b>	Structure moléculaire des protéines de la famille NudC. ....	<b>20</b>
<b>Figure 9:</b>	Isoformes de NudCD1 .....	<b>23</b>
<b>Figure 10:</b>	Niveaux d'expression en ARNm des isoformes de NudCD1 dans différentes lignées cellulaires .....	<b>36</b>
<b>Figure 11:</b>	Niveaux d'expression protéique des isoformes de NudCD1 dans différentes lignées cellulaires .....	<b>37</b>
<b>Figure 12:</b>	Validation des constructions U-2 OS NudCD1 avec étiquette GFP .....	<b>38</b>
<b>Figure 13:</b>	Validation des constructions U-2 OS NudCD1 avec étiquette BirA.....	<b>39</b>
<b>Figure 14:</b>	Dégradation protéasomale des isoformes de NudCD1 .....	<b>40</b>
<b>Figure 15:</b>	Localisation des isoformes de NudCD1 avec l'étiquette GFP en N-terminal.	<b>41</b>
<b>Figure 16:</b>	Localisation des isoformes de NudCD1 avec l'étiquette GFP en C-terminal.	<b>42</b>
<b>Figure 17:</b>	Approches utilisés pour identifier des partenaires d'interaction aux isoformes de NudCD1 .....	<b>43</b>
<b>Figure 18:</b>	Partenaires d'interaction potentiels des isoformes de NudCD1 identifiés par spectrométrie de masse.....	<b>44</b>
<b>Figure 19:</b>	Partenaires d'interaction potentiels des isoformes de NudCD1 identifiés par spectrométrie de masse.....	<b>46</b>
<b>Figure 20:</b>	Regroupement fonctionnel des protéines ayant interagi avec les isoformes de NudCD1-BirA analysées au spectromètre de masse .....	<b>46</b>
<b>Figure 21:</b>	Interaction entre l'isoforme 1 de la protéine NudCD1 et la protéine DHX15 .....	<b>48</b>
<b>Figure 22:</b>	Partenaires d'interaction potentiels de DHX 15 identifiés par spectrométrie de masse .....	<b>49</b>
<b>Figure 23:</b>	Regroupement fonctionnel des protéines ayant interagi avec DHX15 identifiées par spectrométrie de masse .....	<b>50</b>
<b>Figure 24:</b>	Co-localisation de DHX15 avec NudCD1 ou le facteur d'épissage SF3A3 par immunofluorescence dans des cellules U-2 OS .....	<b>51</b>

## LISTE DES ABRÉVIATIONS

4E-BP	eIF4E binding protein
5-FU	5-fluorouracil
a.a.	acide aminé
ADN	acide désoxyribonucléique
ADNc	ADN cyclique
AKT	protein kinase B
Apaf-1	apoptotic protease activating factor 1
Arg	arginine
ARN	acide ribonucléique
ARNm	ARN messenger
ARNPII	ARN polymérase II
ARNsn	petit ARN nucléaire
ARNsh	petit ARN en épingle
ARP1	alpha-centractin
ASK1	apoptosis signal-regulating kinase 1
ATF2	activating transcription factor 2
ATM	ataxia-telangiectasia mutated serine/threonine kinase
ATR	ataxia-telangiectasia and Rad3 related serine/threonine kinase
BAD	Bcl-2 associated death promoter protein
BCR	breakpoint cluster region protein
BirA	bifunctional ligase/repressor BirA
BRAF	RAF homolog B
Brr2	pre-mRNA-splicing helicase Brr2
CBP	cap binding protein
CCT8	T-complex protein 1 subunit theta
CD45	lymphocyte common antigen
cdc25	cell division cycle protein 25
CHERP	calcium homeostasis endoplasmic reticulum protein
c-Kit	proto-oncogene c-kit
CML66	chronic myeloid leukemia 66
COPI	coat protein complex I
CORO1B	coronin-1B
COX-2	cyclooxygenase 2
CTCF	11-zinc finger protein
CTSL	cathepsin L1
CTTN	Src substrate cortactin
Cwc25	pre-mRNA-splicing factor Cwc25
DDB2	DNA damage-binding protein 2
DEAD	sparagine-glutamate-alanine-asparagine
DEAH	asparagine-glutamate-alanine-histidine
DHX15	pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15
DSIF	DRB aensitivity inducing factor
E3	E3 ubiquitin ligase
EEF1D	elongation factor 1-delta
EFHD2	EF-hand domain-containing protein D2



eIF4E	eukaryotic translation initiation factor 4E
Elk-1	ETS domain-containing protein Elk-1
EPB41L3	band 4.1-like protein 3
EPPK1	epiplakin
ERK	extracellular signal-regulated kinase
ESE	exonic splicing enhancer
ESS	exonic splicing silencer
FAS	apoptosis antigen 1
FASN	fatty acid synthase
FECH	ferrochelataase
FLNA	filamin-A
GAPDH	glyceraldehyde 3-phosphate deshydrogenase
GFP	green fluorescent protein
GRB2	growth factor receptor-bound protein 2
GRWD1	glutamate rich WD repeat containing 1
GSK3	glycogen synthase kinase 3
hnRNP	heterogeneous nuclear ribonucleoproteins
HRP	horseradish peroxydase
HSP	heat shock protein
IGF1	insulin-like growth factor 1
IGF1R	IGF1 receptor
I $\kappa$ K	inhibitor of NF- $\kappa$ B kinase
ISE	intronic splicing enhancer
ISS	intronic splicing silencer
JNK	c-jun N-terminal kinase
kDa	kilo dalton
KRAS	kirsten RAS
LEF1	lymphoid enhancer-binding factor 1
LIS1	platelet-activating factor acetylhydrolase 1B subunit 1
LOXL1	lysyl oxidase homolog 1
Lys	lysine
MAP4	microtubule-associated protein 4
MAPK	mitogen-activated protein kinase
MCM	minichromosome maintenance
MDM2	mouse double minute 2 homolog
MED4	mediator of RNA polymerase II transcription subunit 4
MEK	MAPK/ERK kinase
MFSD9	major facilitator superfamily domain-containing protein 9
MKL2	myocardin-like protein 2
MMP15	matrix metallopeptidase 15
MS	spectrométrie de masse
mTORC	mammalian target of rapamycin complex
MUC1	mucin 1
NEDD4	neural precursor cell expressed developmentally down-regulated protein 4
NELF	negative Elongation Factor
NF- $\kappa$ B	nuclear factor-kappa B

Nud	nuclear distribution
NLS	séquence de localisation nucléaire
NOXA	Phorbol-12-myristate-13-acetate-induced protein 1
Ntr	neurotensin receptor type 1
Nud	Nuclear distribution
NudC	Nuclear distribution C
NudCD1	NudC domain containing 1
NudCD1-1	NudCD1 isoforme 1
NudCD1-2	NudCD1 isoforme 2
NudCD1-3	NudCD1 isoforme 3
NudCL	NudC-like
NudCL2	NudC-like 2
OVA66	ovarian cancer-associated antigen 66
p21	cyclin-dependent kinase inhibitor 1
p23	prostaglandin E synthase 3
p38	protein 38 MAPK
p53	tumor protein p53
p150 <sup>glued</sup>	protein 150 glued
PAF1C	RNA polymerase-associated factor 1 complex
PAXBP1	PAX3- and PAX7-binding protein 1
PCPN	PEST proteolytic signal-containing nuclear protein
PCR	réaction de polymérisation en chaîne
PDK-1	phosphoinositide-dependent kinase-1
PDLIM4	PDZ and LIM domain protein 4
PI3K	phosphoinositide 3-kinase
PIK3C2B	phosphatidylinositol 4-phosphate 3-kinase C2 domain-containing subunit beta
PIP	phosphatidylinositol (3,4,5)-triphosphate
Prp	pre-mRNA-processing factor
PSMD2	26S proteasome non-ATPase regulatory subunit 2
P-TEFb	positive transcription elongation factor b
PTEN	phosphatase and tensin homolog
PUMA	p53 upregulated modulator of apoptosis
Pyk2	protein tyrosine kinase 2
qRT-PCR	real-time RT-PCR
RAC1	RAS-related C3 botulinum toxin substrate 1
RAF	rapidly accelerated fibrosarcoma protein
RanBP2	Ran binding protein 2
RAS	rat sarcoma protein
RBM	RNA-binding protein
RCPG	recepteur couplé aux protéines G
RGG	glycing arginine rich
RNA-Seq	Séquençage de l'ARN
RRM	RNA recognition motif domain
RTK	recepteur tyrosine kinase
RT-PCR	reverse transcriptase-PCR
RUVBL	RuvB-like

S100A4	S100 calcium binding protein A4
SB	site de branchement
SC35	Serine/arginine-rich splicing factor SC35
SF1/BBP	splicing factor 1/branch point binding protein
SF2/ASF	splicing factor 2 / alternative splicing factorSF3 splicing factor 3
SHC	Src homology 2 domain-containing
SILAC	stable isotope labeling with amino acids in cell culture
Slu7	pre-mRNA-splicing factor Slu7
Sm	smith
snRNP	petite ribonucléoprotéine nucléaire
Snu114	pre-mRNA-splicing factor Snu114
SOS	son of sevenless
SR	sérine-arginine
SRCRB4D	scavenger receptor cysteine-rich domain-containing group B protein
SRR	domaine de reconnaissance à l'ARN
STIP1	stress-induced-phosphoprotein 1
STK32B	serine/threonine-protein kinase 32B
Sub2	ATP-dependent RNA helicase Sub2
TAGLN2	transgelin-2
TAT-SF1	trans-activator of transcription/steroidogenic factor 1
TFII	transcription factor II
TFIP11	tuftelin-interacting protein 11
Tyr	tyrosine
U	Unité
U2AF	U2 auxiliary factor
U2SURP	U2 snRNP-associated SURP motif-containing protein
uPAR	urokinase receptor
V600E	mutation d'une valine en acide glutamique à la position 600
VEGF	vascular endothelial growth factor
WNT	Wingless-related integration site
WT1	wilms tumor protein
Yju2	protein Cwc16

## REMERCIEMENTS

Je souhaiterais remercier tout d'abord mon directeur de recherche, le Pr François-Michel Boisvert, de m'avoir permis de faire plusieurs stages de recherche ainsi que l'ensemble de ma maîtrise dans son laboratoire malgré une entrevue d'embauche où j'avais littéralement les larmes aux yeux. Il m'a entre autres prodigué des conseils et permis de développer mes connaissances scientifiques, mon esprit analytique ainsi que mes compétences orales et écrites tout au long de mon cheminement. J'aimerais aussi remercier spécialement sa patience par rapport à l'écriture de ce mémoire.

Je veux également remercier très particulièrement les membres du laboratoire, en plus de ceux de l'équipe du Pr Marie-Josée Boucher, que j'ai cotoyés lors de mon temps dans le laboratoire. Dominique Lévesque, pour ses connaissances, son soutien technique, son écoute et ses histoires qui m'ont permis de grandir individuellement et rire à profusion. Marie-Line pour son esprit critique, son support moral et pour toutes les soirées et fins de semaine où elle était là pour rendre le labo vivant et agréable. Un gros merci également à Romain, Alex-Ane, Dominique Jean, Maureen, Mélanie, Benoît, Emanuel et Jennifer qui ont tous été présents et m'ont apporté du support pendant mon cheminement. Un merci aussi à mes amis qui m'ont encouragé, spécialement Benoît Mailhot qui m'a donné de nombreux conseils techniques et Stéphane Handfield pour avoir perdu au jeu. Je salue les professeurs, les collègues étudiants et le personnel que j'ai eu l'opportunité de côtoyer durant ma maîtrise.

Je veux remercier énormément ma famille, Jacqueline, Ralph et Alexandre, pour leur soutien divers et constant tout au long de mes études de la pré-maternelle jusqu'ici. Je tiens à remercier en particulier ma mère pour sa révision linguistique de ce mémoire.

Finalement, merci à mon mentor, le Pr Fernand-Pierre Gendron, ainsi qu'au Pr Alexandre Maréchal et encore une fois à mon superviseur François-Michel, pour leur patience, leur lecture et leurs commentaires sur ce mémoire.

## 1. INTRODUCTION

La cellule humaine est très complexe et se caractérise par un grand nombre de molécules permettant son bon fonctionnement ainsi que celui du corps humain. Les processus biologiques sont principalement orchestrés par des protéines qui interagissent entre elles afin de répondre à des signaux, produire de l'énergie ou proliférer. L'information pour produire ces protéines provient de l'ADN, une longue molécule possédant l'ensemble de l'information génétique dans chaque cellule et conservée lors de chaque réplication cellulaire. Le dogme central de la biologie moléculaire stipule que l'information génétique de l'ADN est transcrite en ARNm avant que celui-ci ne soit traduit en protéine (Crick, 1970). Avant que la protéine soit formée, l'ARNm doit progresser au travers de nombreuses étapes complexes afin de se rendre au cytoplasme sous sa forme finale et d'être reconnu pour sa traduction (Darnell, 2013).

### 1.1 Transcription et épissage

#### 1.1.1 Synthèse du pré-ARNm

Les étapes de la synthèse nécessitent premièrement un remodelage de la chromatine. La chromatine est l'ADN sous forme compacte enroulée sur des octamères d'histones, nommés nucléosomes, permettant la disponibilité ou non-disponibilité de l'ADN, selon le besoin. Le remodelage se fait à l'aide de modifications post-traductionnelles, telles que la méthylation, l'acétylation, l'ubiquitination, la sumoylation ou la phosphorylation, sur des résidus spécifiques au niveau de la partie N-terminale des histones, une partie qui s'étend en forme de queue à la surface des nucléosomes et qui est accessible aux enzymes (Li *et al.*, 2007). Suite à une modification permettant le relâchement de certains nucléosomes, des régions clés du gène se libèrent, telles que le promoteur ou le site d'initiation de l'élongation (Bucceri *et al.*, 2006). La synthèse du pré-ARNm commence avec le recrutement et la formation d'un complexe de préinitiation, composé de plusieurs facteurs généraux de transcriptions, TFIIA, TFIIB, TFIID, TFIIIE, TFIIIF et TFIIH assemblés avec l'ARN polymérase II, ARNP<sub>II</sub>, au promoteur (Orphanides *et al.*, 1996). L'activité hélicase et ATPase de TFIIH est nécessaire pour aider à former la bulle de transcription où l'ADN est

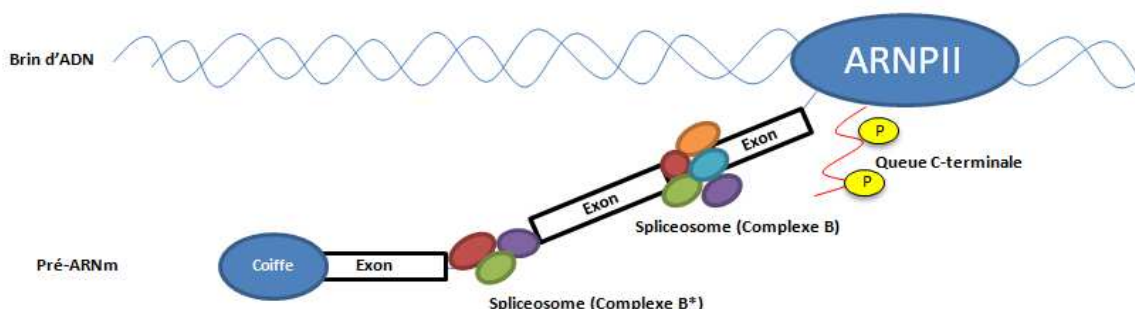
désapparié (Goodrich et Tjian, 1994). L'ARNPII est un complexe protéique, composé de 12 sous-unités, qui transcrit l'ADN en pré-ARNm. L'une de ses sous-unités, RBP1, comporte son domaine c-terminal en forme de queue, composé d'une séquence de 7 a.a. répétée 52 fois. Cette séquence répétée est sujette à plusieurs modifications post-traductionnelles permettant de contrôler l'action de l'ARNPII en tout temps (Egloff et Murphy, 2008). Après transcription de moins de 10 nucléotides, le complexe entre en pause, où la queue de l'ARNPII est d'abord phosphorylée par TFIIF pour permettre le recrutement de NELF et DSIF pour arrêter la transcription immédiatement (Adelman et Lis, 2012).

La transcription peut alors être arrêtée ou bien continuer à l'aide d'une activation par phosphorylation de P-TEFb pour libérer NELF (Nechaev et Adelman, 2011). Le complexe de pré-initiation est alors modifié de manière à laisser certains facteurs généraux sur le promoteur et faciliter les prochaines rondes de transcription (Yudkovsky *et al.*, 2000). De nombreuses autres protéines entrent en jeu dans la régulation de l'élongation par l'ARNPII, principalement suite à une ou des phosphorylations sur les sérines du domaine c-terminal (Lenasi et Barboric, 2013).

Un des facteurs à se lier dans les premiers stades de la transcription est PAF1C. Cette protéine permet, en plus de l'élongation, des modifications sur les histones avoisinantes et aide également dans le traitement de l'ARNm (Zheng *et al.*, 2010). Lorsque le transcrit de pré-ARNm atteint environ 20 nucléotides, une coiffe de 7-méthylguanosine est formée à la suite d'une série de réactions catalysées par 3 enzymes au niveau de l'extrémité 5' du transcrit pré-ARNm (Jurado *et al.*, 2014). Les protéines CBP20 et CBP80 lient alors la coiffe afin de former un complexe impliqué dans l'épissage (Izaurralde *et al.*, 1994). La fin de l'élongation nécessite la reconnaissance d'une séquence commune reconnue près de la fin du transcrit d'ARNm, AAUAAA (Proudfoot, 2011). Des complexes de clivage et de polyadénylation reconnaissent cette séquence sur l'ARNm afin de cliver celui-ci du complexe transcriptionnel, ainsi que de faire la synthèse d'une queue polyA sur la partie 3' du transcrit clivé (Proudfoot, 2011).

### 1.1.2 Épissage du pré-ARNm

Parallèlement à sa synthèse, le pré-ARNm est traité de manière à devenir un ARNm (Figure 1). Un pré-ARNm est un transcrit d'ARN composé de séquences codantes, les exons, et de séquences non-codantes, les introns, qui sont liées ensemble. Le traitement qui permet au pré-ARNm de devenir un ARNm, dépourvu d'introns, se nomme épissage (Breathnach et Chambon, 1981). Ce mécanisme est catalysé par un complexe ribonucléoprotéique nommé spliceosome (Fredericks *et al.*, 2015). Le spliceosome est un large complexe contenant 5 snRNP, soit U1, U2, U4/U6 et U5, ainsi que de nombreuses autres protéines (Zhou *et al.*, 2002). Chaque snRNP est constituée d'un ou deux snARN uniques, un ensemble commun de 7 protéines Sm, et d'un nombre variable d'autres protéines (Stark et Lührmann, 2006).

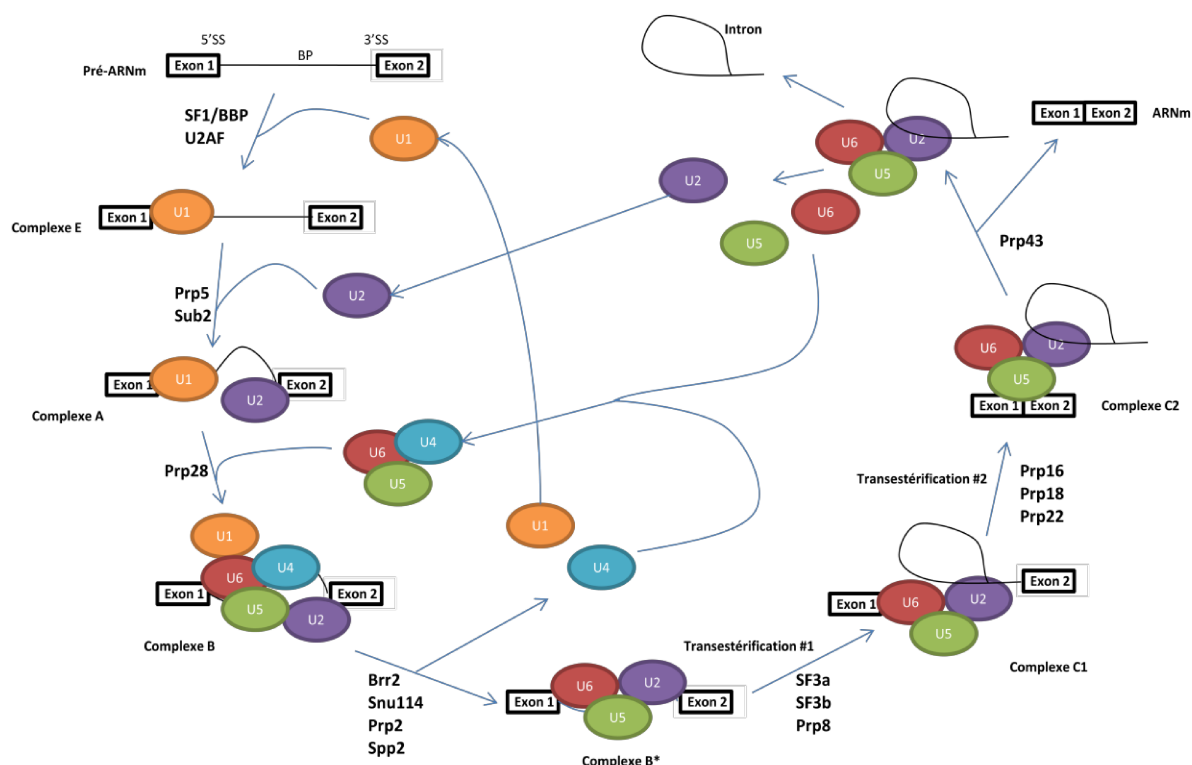


**Figure 1: Synthèse et modifications du pré-ARNm.** Les modifications comme l'ajout de coiffe et l'épissage alternatif du pré-ARNm s'effectuent alors que le brin est toujours en phase d'élongation. Ces différents mécanismes sont interreliés et travaillent conjointement et simultanément à la synthèse de l'ARNm.

L'assemblage du spliceosome autour de l'intron qui doit être épissé au niveau du pré-ARNm nécessite d'abord la présence de séquences conservées au point de coupure en 5', en 3', une séquence riche en pyrimidines et au site de branchement (SB) de l'intron (Smith *et al.*, 1993). L'assemblage du spliceosome débute par la formation du Complexe E' (Kent *et al.*, 2004). Le snRNP U1 reconnaît premièrement le site 5' avant de s'y lier, qui permet donc à la protéine SF1/BBP d'aller lier le SB (Kent *et al.*, 2004). La protéine hétérodimérique U2AF vient alors lier SF1/BBP avec sa sous-unité 65 pour ensuite lier la

séquence riche en pyrimidines et permettre à sa sous-unité 35 de reconnaître le site 3' et ainsi former le complexe E (Figure 2) (Zhang *et al.*, 2013). Subséquemment, le snRNP U2, à l'aide des ARN hélicases Prp5 et Sub2, déplace SF1/BBP pour se lier au SB de manière ATP-dépendante pour ainsi former le complexe A (Figure 2) (Wahl *et al.*, 2009). Les snRNP U4/U6 et U5, pré-assemblées en complexe, sont recrutées au spliceosome à l'aide de l'hélicase Prp28 de manière ATP-dépendante pour former le complexe B (Matera et Wang, 2014). Ce complexe est modifié par l'action de Brr2, Snu114, Prp2 et Spp2, qui mène à la libération de U4 et U1 du complexe B pour former le complexe B\*, catalytiquement actif (Figure 2) (Silvermann *et al.*, 2004, Matera et Wang, 2014). Certaines de ces protéines sont principalement caractérisées chez la levures. Ces protéines possèdent toutefois des homologues conservées chez l'humain, tel que SNRNP20 pour Brr2 et EFTUD2 pour Snu114 (van der Feltz *et al.*, 2012).

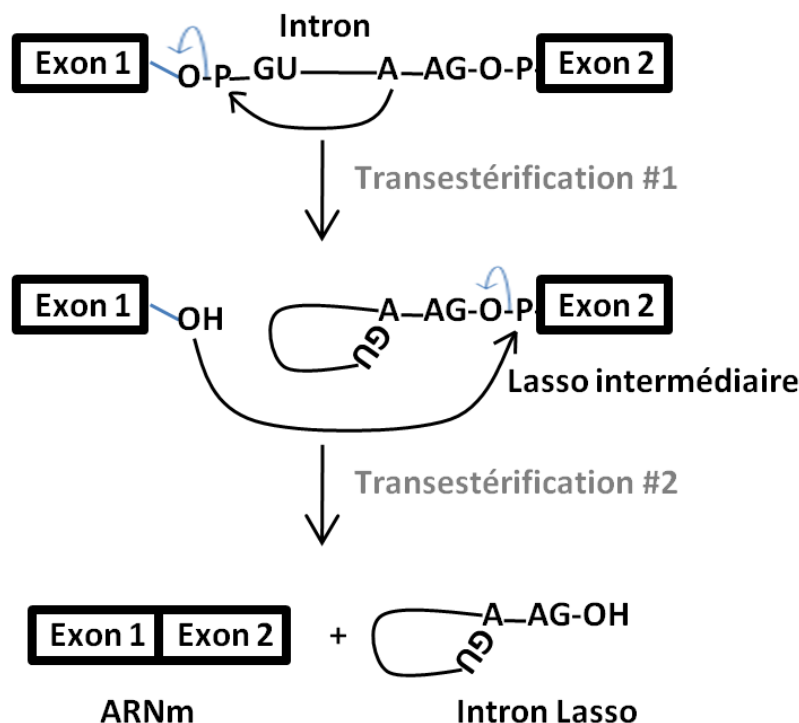




**Figure 2: Épissage du pré-ARNm par le spliceosome.** Représentation schématisée des différentes étapes de la formation des complexes du spliceosome sur le pré-ARNm. La formation du complexe débute par l'ajout du snRNP U1 sur le pré-ARNm pour former le complexe E. Le snRNP U2 s'insère par la suite pour former le complexe A. Le complexe B est créé suite à l'insertion du tri-snRNP formé de U4 U5 et U6. La réaction catalytique de l'épissage s'effectue suite à la formation du complexe B\*, impliquant un changement de conformation et une perte de U1 et U4, jusqu'au complexe C2. Le désassemblage permet la libération de l'ARNm et des protéines associées, de l'intron et des snRNP qui sont recyclés pour participer à un nouveau cycle d'épissage. Seules les snRNP ainsi que certaines protéines clés ont été schématisées sur cette image.

La première étape de l'épissage de l'intron survient grâce aux protéines SF3a et SF3b, qui exposent le SB, et de Prp8 aidée par Cwc25 et Yju2 pour permettre une attaque nucléophile entre le groupement 2'-OH du SB et le site d'épissage 5', formant une structure en forme de lasso au niveau de l'intron et laissant l'extrémité de l'exon en 5' libre, ce spliceosome est alors sous forme de complexe C1 (Figure 3) (Lardelli *et al.*, 2010, Horowitz, 2012). Après des réarrangements ATP-dépendants de Prp8, Prp16, Prp18, Prp22, Ntr1, Ntr2 et Slu7 sur la conformation du complexe, cette extrémité libre attaque alors le site d'épissage 3', et mène à la liaison des exons en 5' et en 3' de l'intron retiré pour former le complexe C2 (Figure 3) (Horowitz, 2012). Après avoir effectué ses actions catalytiques, le spliceosome

doit être désassemblé afin de relâcher l'ARNm, les snRNP restantes, ainsi que les facteurs d'épissages (Figure 2) (Fourmann *et al.*, 2013). L'activité de Prp22 permet de déstabiliser l'interaction entre l'intron et l'ARNm pour libérer l'ARNm du spliceosome (Figure 2) (Schwer, 2008). Ntr1 et Ntr2 recrutent Prp43 au spliceosome afin de libérer les snRNP U2, U5 et U6 d'entre eux et de l'intron clivé et de permettre leur recyclage pour d'autres rondes d'épissage (Figure 2) (Fourmann *et al.*, 2013).

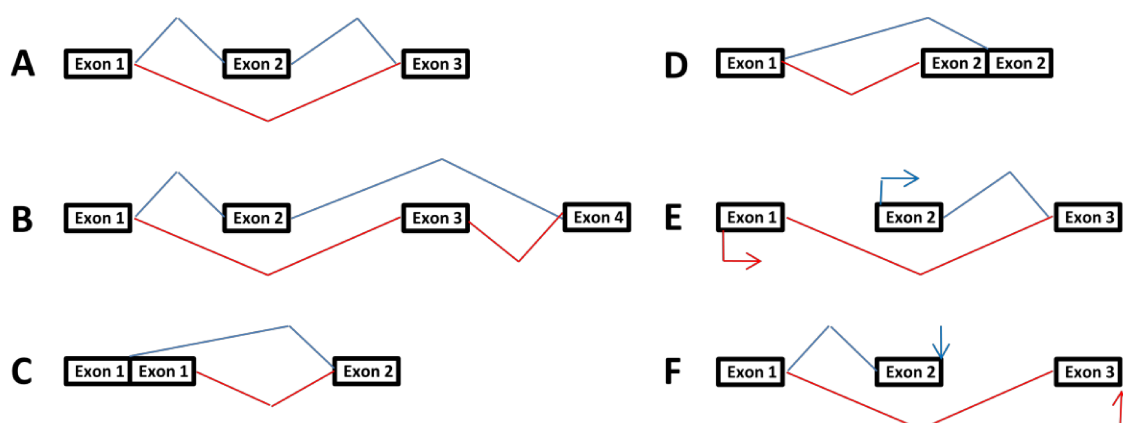


**Figure 3: Transestérification du complexe C.** Les deux étapes de transestérification catalysées pour enlever l'intron lors de l'épissage.

### 1.1.3 Épissage alternatif du pré-ARNm

L'épissage alternatif est un mécanisme de maturation du pré-ARNm permettant l'obtention d'une plus grande diversité protéique malgré un nombre relativement limité de gènes en incluant ou non certains exons dans l'ARNm final (David et Manley, 2008). Il a été révélé que plus de 90% des 24 000 gènes chez l'humain subissaient de l'épissage alternatif, permettant aux 100 000 différentes protéines humaines d'être synthétisées (Keren *et al.*, 2010). Les protéines issues d'un même gène, mais épissées alternativement sont nommées

isoformes. Ces différences d'exons entre les isoformes peuvent mener à des changements de localisation, d'interaction protéique, de site de modification post-traductionnel et à d'autres divergences chez ces protéines. Un pré-ARNm possédant plusieurs exons peut être épissé de différentes manières. Bien que la plupart des exons d'un pré-ARNm sont constitutifs et sont inclus dans l'ARNm final, il existe des exons qui sont parfois inclus ou exclus dans l'ARNm, nommés des exons cassettes (Figure 4A) (Black, 2003). Certains pré-ARNm peuvent posséder plusieurs de ces exons cassettes et ceux-ci peuvent être mutuellement exclusifs, produisant donc des ARNm possédant l'un ou l'autre de ces exons (Figure 4B) (Smith et Nadal-Ginard, 1989). Les exons peuvent également être épissés sur des sites d'épissage 5' ou 3' alternatifs, pouvant ainsi augmenter ou diminuer la taille de l'exon (Figure 4C, Figure 4D) (Black, 2003). Ces coupures alternatives en 5' peuvent aussi être contrôlées par l'action de promoteurs alternatifs, alors que celles en 3' peuvent être causées par des sites de polyadénylation alternatifs (Figure 4E, Figure 4F) (Black, 2003). Ces différentes manières d'épissage combinées sous l'actions de divers facteurs peuvent donc permettre à un pré-ARNm d'être épissé en divers ARNm.



**Figure 4: Types d'épissage alternatif.** Schématisation de l'épissage alternatif d'exons cassettes (A), d'exons cassettes mutuellement exclusifs (B), d'épissage de 5' (C) ou 3' (D) alternatifs, d'épissage régi par promoteurs alternatifs (E) et par sites de polyadénylation alternatifs (F).

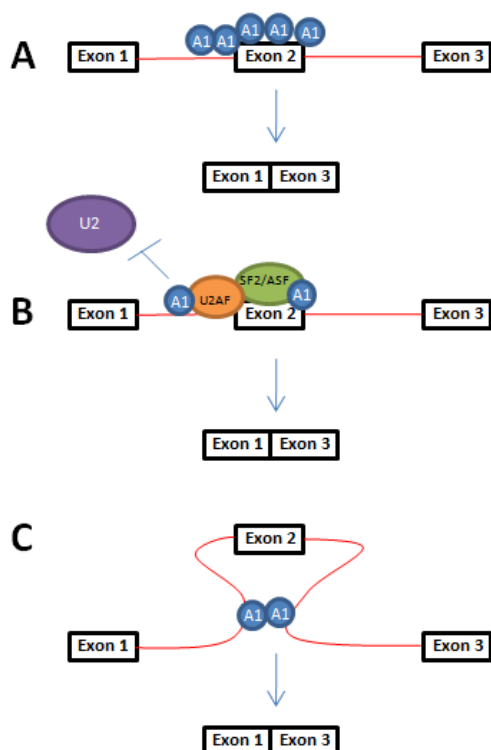
#### 1.1.4 Régulation de l'épissage

La régulation qui dicte les exons inclus et les exons exclus du transcrit d'ARNm final implique des éléments de séquence ARN ainsi que des protéines régulatrices (Chen et

Manley, 2009). Les éléments de séquence sont divisés en quatre catégories fonctionnelles : les *exonic splicing enhancers* (ESE), les *exonic splicing silencers* (ESS), les *intronic splicing enhancers* (ISE) et les *intronic splicing silencers* (ISS) (Chen et Manley, 2009). Les ESE sont principalement reconnus et liés par des membres de la famille de protéines sérine-arginine (SR) (Long et Caceres, 2009). Les protéines SR forment une famille de facteurs nucléaires qui sont impliqués dans la régulation de par leur capacité à lier les ESE via leurs domaines de reconnaissance à l'ARN (SRR), et de médier les interactions protéiques grâce à leur domaine SR riche (Kohtz *et al.*, 1994). Le domaine SR riche sert également de NLS chez cette famille de protéines (Lai *et al.*, 2001). Le fonctionnement principal de ces protéines est de faciliter la reconnaissance d'un site d'épissage particulier en favorisant le recrutement d'U2AF et U1 sur celui-ci avec son domaine SR, après reconnaissance et liaison du ESE par leur domaine SRR (Bourgeois *et al.*, 1999). Les protéines SR peuvent également interagir entre elles pour former de plus grands complexes favorisant l'épissage (Longman *et al.*, 2001).

Une autre famille de protéine est prédominante dans la régulation de l'épissage: les protéines hnRNP, qui lient normalement les séquences d'inhibition ESS et ISS (Chen et Manley, 2009). Celles-ci sont constituées de plusieurs motifs de liaison à l'ARN, principalement les motifs RRM, permettant la liaison à l'ARN, et de la boîte RGG, responsable des interactions homologues ou hétérologues avec les autres hnRNP (Geuens *et al.*, 2016). Cette famille possède 16 membres ayant leurs particularités en termes de domaines et de motifs de liaison à l'ARN (Geuens *et al.*, 2016). La plus étudiée de cette famille est hnRNP A1. Elle peut agir de 3 manières différentes au niveau des ESS et ISS pour mener à l'excision des exons et introns spécifiques (Black, 2003). La première méthode implique son association à un site silencer et recrute d'avantage de hnRNP A1 pour créer une zone de blocage à l'association du spliceosome de U2AF (Figure 5A) (Zhu *et al.*, 2001). La deuxième méthode implique son association à un site adjacent au SB, n'empêchant pas la venue de U2AF, mais plutôt celle d'U2 au SB (Figure 5B) (Tange *et al.*, 2001). La troisième méthode implique la liaison de hnRNP A1 envers des sites d'associations introniques de différents introns et d'une multimérisation permettant de créer une boucle contenant l'exon, qui sera donc exclue du transcrit final (Figure 5C) (Blanchette et Chabot, 1999). Ces processus peuvent également être influencés par les protéines SR,

qui ont un rôle inhibiteur sur l'action des hnRNP, alors que leur liaison à un ESE près d'un ESS où un hnRNP y est attaché causerait un échec du mécanisme d'action de hnRNP (Long et Caceres, 2009, Tange *et al.*, 2001). Les ISE sont des éléments moins caractérisés que les ESE. Certaines études montrent toutefois que ceux-ci pourraient nécessiter l'interaction avec des protéines SR, de hnRNP et même d'autres protéines (Hastings *et al.*, 2001, Black, 2003).



**Figure 5: Modèles de répression de l'épissage par hnRNP A1.** L'hnRNP A1 peut influencer l'épissage de différentes manières. Elle peut s'associer à un site silencier et recruter d'autres A1 pour bloquer la formation du spliceosome (A). Elle peut également empêcher l'association du snRNP U2 au SB (B). A1 peut finalement reconnaître des sites de liaisons introniques, s'y lier et multimériser pour empêcher la reconnaissance de l'exon par le spliceosome (C).

### 1.1.5 Relations transcription et épissage

La transcription et l'épissage sont deux processus cellulaires intimement liés. Ceux-ci ne se font pas de manière indépendante, mais bien en s'influençant de part et d'autre. Un

exemple serait l'action de WT1, un facteur de transcription, qui lie également U2AF65 pour promouvoir l'épissage. La partie C-terminale peut influencer l'épissage de diverses manières. Sa phosphorylation aide à l'assemblage des complexes du spliceosome alors qu'inversement, sa non-phosphorylation inhibe la formation des complexes (Hirose *et al.*, 1999). Son domaine C-terminal peut également aider à la reconnaissance d'exons pour faciliter le positionnement lors de l'épissage (Zeng et Berget, 2000). La structure du promoteur peut également influencer le recrutement de facteurs de transcription sur l'ARNPII, permettant alors leurs interactions avec des facteurs d'épissage (Kornblihtt *et al.*, 2004).

Une importante quantité de protéines du spliceosome n'est pas directement impliquée dans l'épissage, mais plutôt via d'autres rôles dans l'expression génique (Zhou *et al.*, 2002). Un exemple de communication impliquant des facteurs d'épissage et l'ARNPII est l'association de SF2/ASF, un membre de la famille de protéines SR, avec le domaine C-terminal de l'ARNPII ayant des sérines-2 phosphorylées qui permettrait son interaction avec les sites ESE sur le pré-ARNm pour réguler l'épissage alternatif (Barboric, 2009, Kornblihtt *et al.*, 2004, Das *et al.*, 2007). Le snRNP U1, faisant partie du spliceosome, a également un contrôle sur la transcription en facilitant la réinitialisation de transcription médiée par TFIIH (Kwek *et al.*, 2002). U1 peut également interagir avec TAT-SF1, un facteur d'élongation de la transcription pour stimuler l'élongation de l'ARNPII (Fong et Zhou, 2001). Le complexe médiateur peut également influencer l'épissage, via sa sous-unité MED23 qui interagit spécifiquement avec hnRNP L lors de l'épissage alternatif de certains exons (Huang *et al.*, 2012). Une étude a montré que la phosphorylation par P-TFEB sur la Ser2 de la queue de l'ARNPII promeut l'inclusion de l'exon EDA en impliquant le recrutement de SF2/ASF, alors qu'une phosphorylation insuffisante de ces résidus mène à l'exclusion de cet exon (Barboric, 2009). La vitesse de transcription possède également un impact sur l'épissage. Une étude a démontré une densité supérieure d'ARNPII sur l'ADN lors de la transcription d'exons, permettant plus de temps aux facteurs d'épissage de reconnaître les différents sites où ils se lient (Chodavarapu *et al.*, 2010). Les phosphorylations sur le domaine C terminal de l'ARNPII mènent à des augmentations de vitesse de transcription pouvant ainsi causer l'exclusion d'un plus petit exon ayant des sites ESE plus faibles. Cette exclusion résulte de l'assemblage du

spliceosome à un point d'épissage plus fort en aval en raison d'une non-reconnaissance causée par la transcription trop rapide, causant donc le saut de l'exon. Inversement, une transcription plus lente permet aux protéines SR plus de temps pour s'insérer au site d'épissage et d'influencer la formation du spliceosome malgré le site ESE plus faible. Les différentes phosphorylations de l'ARNPII peuvent donc augmenter ou diminuer la vitesse de transcription et ainsi avoir un effet sur l'épissage des pré-ARNm. Ce même mécanisme peut également impliquer de l'épigénétisme, avec l'exemple de CTCF, une protéine avec domaine de liaison à l'ADN (Shukla *et al.*, 2011). L'action de CTCF permet une pause de l'élongation de l'ARNPII, favorisant inclusion de l'exon 5 du gène de CD45 dans l'ARNm (Shukla *et al.*, 2011). La liaison de CTCF est inhibée lors de méthylation de l'ADN, empêchant la pause et contribuant à l'exclusion de l'exon 5 de CD45 (Shukla *et al.*, 2011). Les pauses permettent également de faciliter l'ajout de coiffe en 5' et de la synthèse de la queue polyA en 3' de l'ARNm (Perales et Bentley, 2009). Les enzymes de coiffe sont par ailleurs réimpliquées au niveau de la transcription (Mandal *et al.*, 2004). Ces enzymes permettent, au travers d'interactions protéiques, de rétablir l'élongation après l'arrêt médié par NELF, avant de favoriser le recrutement de P-TEFb (Mandal *et al.*, 2004).

## 1.2. Cancer

Le cancer est la principale cause de décès au Canada, comptant pour presque 30% des décès (Canada, 2016). Un cancer est un tissu complexe composé de plusieurs types cellulaires qui s'influencent afin de favoriser la croissance et la progression tumorale (Hanahan et Weinberg, 2011). Dix caractéristiques particulières ont été identifiées chez les cellules tumorales malignes. Elles ont la capacité de proliférer de façon soutenue, elles sont insensibles aux signaux d'arrêt de prolifération provenant des protéines suppresseurs de tumeurs, elles sont résistantes à la destruction immunitaire, elles ont une immortalité répllicative, elles favorisent la formation d'un environnement inflammatoire favorable, elles sont invasives et métastatiques, elles induisent l'angiogénèse, elles mutent constamment, elles ont une résistance aux mécanismes de mort cellulaire programmée et elles ont un métabolisme altéré (Hanahan et Weinberg, 2011).

La protéine la plus souvent au coeur de la formation tumorale est la protéine p53, qui est mutée à différents niveaux dans plus de 50% des cancers (Lane et Levine, 2010). Cette protéine est un facteur de transcription qui contient différents domaines permettant sa tétramérisation, sa liaison à l'ADN, sa localisation ainsi que pour réguler ses modifications post-transcriptionnelles (Dai et Gu, 2010). p53 est hautement régulée par la protéine MDM2, une E3 ubiquitin ligase qui cause principalement une dégradation constante de p53 en polyubiquitinant son domaine de modifications post-traductionnelles, permettant de garder son niveau d'expression très faible chez les cellules saines (Brooks et Gu, 2006). Chez les cellules saines, des phénomènes comme des dommages à l'ADN peuvent causer l'activation de kinases telles qu'ATM et ATR qui phosphorylent et ainsi inactivent MDM2, empêchant la dégradation de p53 (Meek, 2015). Ainsi p53, lorsqu'à un niveau d'expression suffisant, permet différents phénomènes cellulaires anti-tumoraux comme l'arrêt de croissance, qui permet entre autres la réparation de l'ADN, la sénescence, l'apoptose, ou même l'inhibition de l'angiogénèse (Vousden et Lane, 2007, Teodoro *et al.*, 2007). Les mutations de p53 peuvent aller jusqu'à causer des problèmes de liaisons à l'ADN, dérangeant ainsi son activité dans la transcription, lui permettant même souvent de gagner de nouvelles fonctions permettant de promouvoir des caractéristiques tumorales comme l'invasion, la prolifération ou l'instabilité génomique chez ces cellules (Muller et Vousden, 2014, Meek, 2015). p53 régule les phénomènes anti-tumoraux en se liant à divers gènes, incluant ceux régulant p21 et CDC25C pour initier un arrêt du cycle cellulaire, d'autres gènes comme ceux régulant PUMA, NOXA et APAF-1 pour mener à l'apoptose, ou même DDB2, impliquée dans la réparation des dommages à l'ADN (Riley *et al.*, 2008, Lane et Levine, 2010).

Outre l'action de p53, plusieurs autres voies de signalisations qui sont impliquées dans la réplication ou la mort cellulaire normale des cellules sont modifiées chez des cellules cancéreuses. Une de ces voies de signalisation est la signalisation par le récepteur IGF1R, activant notamment les voies de signalisation subséquentes de PI3K/AKT et des MAPK (Haisa, 2013).

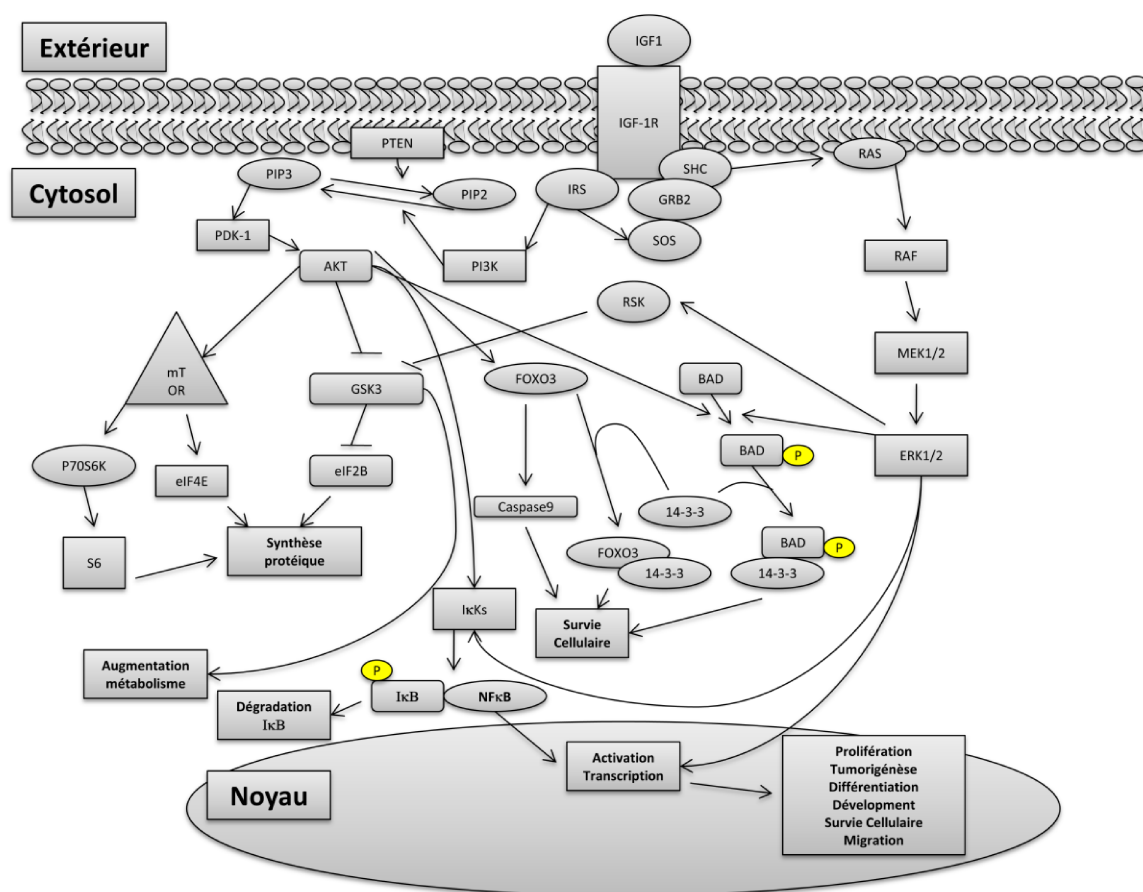


### 1.2.1 Voie de signalisation IGF1R

La voie IGF1R est démontrée comme surexprimée ou constitutivement active dans un grand nombre de lignées cellulaires en raison de sa cascade menant à des effets comme la différenciation, la transformation et la prévention de l'apoptose (Fürstenberger et Senn, 2003). La signalisation IGF1R (Figure 6) est une voie impliquée dans divers processus de croissance lors du développement et continue ses effets chez l'individu adulte. Ce récepteur à tyrosine kinase (RTK) est constitué de deux chaînes  $\alpha$  extra-cellulaires, pour interaction avec le ligand, ainsi que de deux chaînes  $\beta$  transmembranaires comprenant un domaine tyrosine kinase cytoplasmique et un domaine c-terminal (Massagué et Czech, 1982). Le système d'IGF1R est normalement activé par les ligands IGF1, IGF2 et l'insuline (Ullrich *et al.*, 1986). Ces ligands lient des domaines des chaînes  $\alpha$  pour transmettre un signal vers les chaînes  $\beta$  transmembranaires (Sundaram, 2013). Ce signal cause des changements conformationnels stimulant l'activité tyrosine kinase, suivie par une autophosphorylation des Tyr1131, 1135 et 1136 de la portion intermembranaire du récepteur (Li *et al.*, 2009). Cette autophosphorylation augmente l'activité tyrosine kinase d'IGF1R, qui phosphoryle alors d'autres sites de la portion intermembranaire ainsi que des protéines venant s'y attacher (Kuemmerle, 2002). Deux protéines phosphorylées particulièrement par IGF1R sont IRS1 et IRS2 (Copps et White, 2012). Leur phosphorylation permet le recrutement, via le domaine SH2, d'une sous-unité catalytique de PI3K, membre-clé de la voie PI3K/AKT (Wybenga-Groot et McGlade, 2015). L'autophosphorylation d'IGF1R cause également le recrutement et l'activation de la protéine SHC (Perrini *et al.*, 2010). Celle-ci lie l'adaptateur GRB2 qui recrute et active SOS (Perrini *et al.*, 2010). Cette activation mène à une autre activation, soit celle de membres de la famille de petites protéines G, RAS, et subséquemment de la voie de signalisation des MAPK (Chen *et al.*, 1996). IRS1 et IRS2 peuvent également activer la voie de signalisation des MAPK en recrutant et activant GRB2 et SOS (Copps et White, 2012).

Le domaine c-terminal est notamment impliqué dans ses modifications post-traductionnelles et sa signalisation (Haisa, 2013). MDM2, pouvant également être cytosolique, connue comme ayant un rôle de régulation sur p53, a été identifiée comme interagissant avec IGF1R dans un processus d'ubiquitination menant à la régulation

négative des signaux du récepteur, via une augmentation de sa dégradation (Girnita *et al.*, 2003). MDM2 n'interagit pas directement avec IGF1R, mais grâce à la protéine  $\beta$ -arrestin1 (Girnita *et al.*, 2005). Ces protéines mènent toutefois à une troisième manière d'activer la voie des MAPK, découverte plus récemment (Girnita *et al.*, 2007). Lors de l'activation d'IGF1R, MDM2 peut ubiquitinyler  $\beta$ -arrestin1 pour permettre à celle-ci de jouer un rôle dans l'activation de la voie des MAPK (Girnita *et al.*, 2007). Une autre E3 ubiquitin ligase, NEDD4, est démontrée comme interagissant avec GRB10, une autre protéine adaptatrice impliquée lors de l'activation d'IGF1R qui régule négativement la prolifération médiée par IGF1R (Monami *et al.*, 2008). De récentes études démontrent toutefois qu'en plus d'un rôle dans l'ubiquitination, l'internalisation et la dégradation du récepteur, l'internalisation par la E3 ligase NEDD4 serait également impliquée dans la régulation positive de la voie PI3K/AKT (Monami *et al.*, 2008, Zhang *et al.*, 2015). Une autre voie d'activation d'IGF1R passerait par sa sumoylation par RanBP2 (Packham *et al.*, 2015). Cette sumoylation permet la translocation d'IGF1R jusqu'au noyau qui permettrait d'activer la transcription de gènes en s'associant au facteur de transcription LEF1, causant la traduction de gènes tels que CYCLIN D1 et AXIN2, impliqués dans la régulation du cycle cellulaire et de la voie WNT, favorisant la prolifération cellulaire (Warsito *et al.*, 2012)



**Figure 6: Voie de signalisation IGF1R.** Représentation graphique simplifiée de la voie de signalisation IGF1R. L'activation du RTK IGF1R par ses ligands permet l'autophosphorylation du récepteur et l'activation subséquente des voies de signalisation des MAPK et d'AKT. Ces voies de signalisation activent alors plusieurs partenaires protéiques cytosoliques et nucléaires en vue de favoriser des phénomènes de prolifération, différenciation, synthèse protéique, survie cellulaire et autres caractéristiques cellulaires.

### 1.2.2 Voie de signalisation AKT

La voie de signalisation AKT commence tout d'abord par la phosphorylation de PI3K pour catalyser la conversion de PIP2 (inactif) en PIP3 (actif) (Kuemmerle, 2002). La réaction inverse pour convertir le PIP3 en PIP2 est catalysée par la protéine PTEN (Georgescu, 2010). Le PIP3 permet le recrutement de PDK-1 et AKT à la membrane, donnant lieu à la phosphorylation subséquente d'AKT par PDK-1 (Wick *et al.*, 2000). La kinase AKT phosphorylée est activée et phosphoryle à son tour un grand nombre de protéines afin d'activer ou d'inhiber celles-ci. AKT activée phosphoryle le complexe mTORC1, celui-ci

entre autres phosphorylant 4E-BP pour permettre la libération de eIF4E et ainsi stimuler la synthèse protéique (Martini *et al.*, 2014). Le complexe mTORC2 peut également contribuer à la phosphorylation et l'activation complète d'AKT via sa phosphorylation sur la sérine 473 (Martini *et al.*, 2014). La protéine AKT activée va également influencer la signalisation par le facteur de transcription NF- $\kappa$ B via les kinases I $\kappa$ K. Celles-ci phosphorylent les I $\kappa$ B, qui lient NF- $\kappa$ B pour l'empêcher de migrer au noyau. Les I $\kappa$ B phosphorylées sont dégradées et permettent à NF- $\kappa$ B d'activer la transcription de gènes anti-apoptotiques (Huang *et al.*, 2015). Des protéines pro-apoptotiques de la famille de facteur de transcription forkhead sont également visées par AKT, notamment par sa phosphorylation de la protéine FOXO3, empêchant l'activité de celle-ci (Zheng *et al.*, 2002). Plusieurs protéines de cette voie de signalisation peuvent contribuer à une suractivation de cette voie et mener à un phénotype cancéreux, par exemple par la perte de PTEN, une mutation activatrice d'AKT ou une mutation activatrice de PI3K, qui sont retrouvées dans différentes instances de cancer du sein (Deng *et al.*, 2015).

### 1.2.3 Voie de signalisation des MAPK

Le récepteur IGF1R peut également activer la voie de signalisation passant par les MAPK. La voie passant par les MAPK peut activer 4 sous-voies ayant en commun 2 sérine/thréonine kinases et une thréonine/tyrosine kinase, la voie de MAPK/ERK, la voie MAPK7, la voie JNK et la voie p38 (Dhanasekaran et Premkumar Reddy, 1998, Robinson et Cobb, 1997). Les voies JNK et p38 sont induites en réponse à des stress cellulaires et peuvent mener à l'activation de gènes anti et pro-apoptotiques (Wada et Penninger, 2004). La voie canoniquement activée est celle de MAPK/ERK (Morrison, 2012). Cette voie est initialement activée par un récepteur couplé aux protéines G (RCPG) ou un RTK. Les RCPG sont un type de récepteur composé de 7 domaines transmembranaires, qui, lorsqu'activé par un ligand, activent la phosphorylation de protéines hétérotrimériques par du GTP, menant à une cascade d'activation en aval. Suite à l'activation des RAS, celles-ci activent en cascade les MAPK kinase kinases, RAF, qui activent à leur tour en aval les MAPK kinases MEK1 et MEK2, qui finalement phosphorylent les MAPK ERK1 et ERK2 (Morrison, 2012). L'activation des ERK1/2 a tout d'abord un effet au niveau du cytoplasme en phosphorylant différentes cibles, comme GSK3 influençant le métabolisme, BAD

influençant la survie cellulaire et d'autres protéines participant à des activités comme le mouvement et l'adhésion cellulaire (Pullikuth et Catling, 2007, Park *et al.*, 2013, Sun *et al.*, 2015). ERK1/2 transloque ensuite vers le noyau où il active plusieurs facteurs de transcription liés à des processus comme la prolifération, la différenciation, le développement et la survie cellulaires (Morrison, 2012, Burroto *et al.*, 2014). Des mutations, comme la mutation V600E chez BRAF et multiples mutations repertoriées chez KRAS, mènent à une suractivation de la voie des MAPK et à la tumorigénèse chez ces cellules (Burroto *et al.*, 2014).

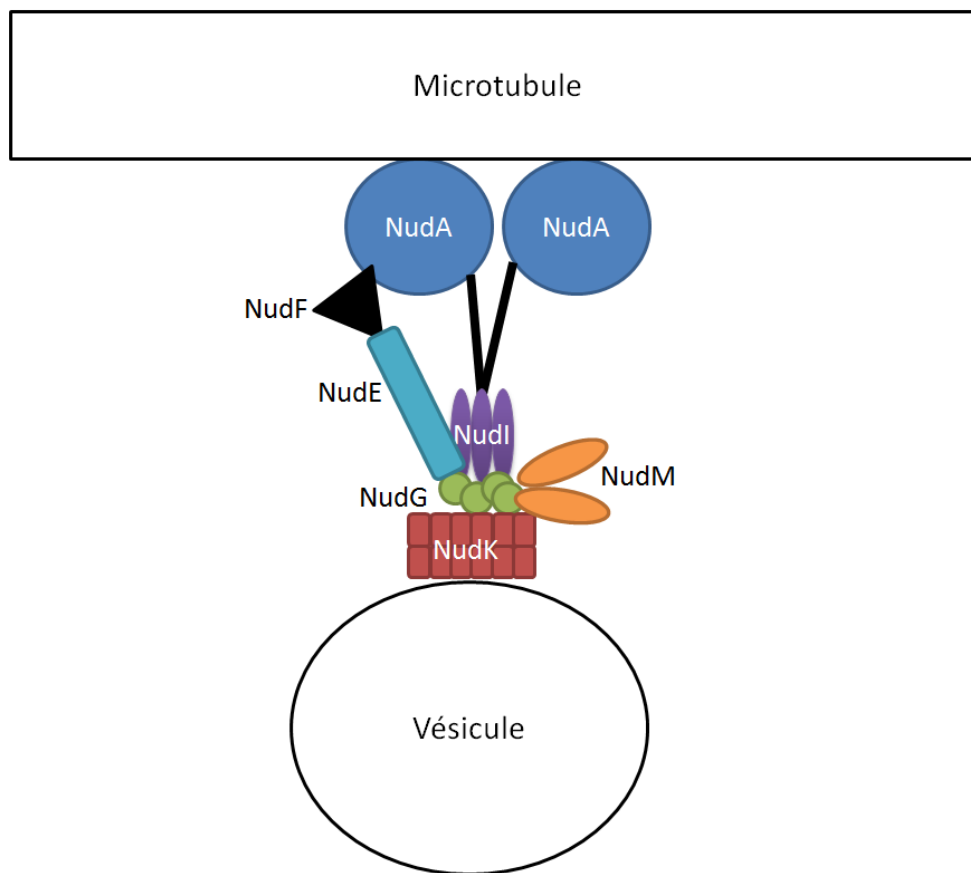
### 1.3. NudCD1

#### 1.3.1 Famille Nud

Afin d'assurer un bon développement et une bonne croissance au niveau des cellules, des mécanismes de transport vésiculaire et de migration nucléaire sont nécessaires. Ces mécanismes sont en partie contrôlés par des protéines nommées kinésines, dynéines cytoplasmiques et dynactines (Walker et Sheetz, 1993). Ces protéines sont en effet impliquées dans divers processus de transport microtubule-dépendants nécessaires à la survie, la prolifération et la croissance cellulaire (Walker et Sheetz, 1993). Les membres de cette famille sont nommés ainsi selon leurs composantes identifiées chez l'organisme *A. nidulans*. Plusieurs des gènes de la famille Nud, acronyme de *nuclear distribution*, codent pour des sous-unités de la dynéine ou de la dynactine (Figure 7) (Morris, 2000). La dynéine est un complexe protéique ATP-dépendant formé de deux chaînes lourdes, deux à trois chaînes intermédiaires et au minimum quatre chaînes légères (Walker et Sheetz, 1993). La dynactine pour sa part est également un complexe protéique nécessaire au transport vésiculaire médié par la dynéine et la kinésine en servant d'attachement entre la vésicule et la dynéine ou kinésine (Bruno *et al.*, 1996). Plus particulièrement, NudA, NudI et NudG composent respectivement les chaînes lourdes, intermédiaires et légères de la dynéine cytoplasmique, alors que pour leur part, NudK et NudM codent pour des composantes importantes de la dynactine, connues sous le nom d'ARP1 et p150<sup>glued</sup> chez les humains (Schroer, 1994, Morris, 2000, Riera et Lazo, 2009). Chez les cellules différenciées, la dynéine cytoplasmique est impliquée dans le transport rétrograde d'organelle le long des

microtubules, dans l'endocytose, l'organisation de l'appareil de Golgi, ainsi que le transport axonal rétrograde (Walker et Sheetz, 1993). Le transport rétrograde contrôlé par ces protéines inclut entre-autres le transport d'endosomes, lysosomes, phagosomes, peroxysomes et de vésicules du réticulum endoplasmique vers l'appareil de Golgi (Roberts *et al.*, 2013). Les dynéines cytoplasmiques permettent aussi la formation des fuseaux mitotiques lors de la division cellulaire en s'attachant aux structures cellulaires comme la membrane cytoplasmique pour tirer les deux fuseaux aux deux extrémités cellulaires et aident également lors de mécanismes moléculaires permettant la bonne ségrégation des chromosomes lors de la métaphase en retardant l'anaphase jusqu'à ce que ceux-ci soient correctement attachés aux fuseaux (Roberts *et al.*, 2013).

Il y a toutefois 3 gènes de cette famille qui codent pour d'autres protéines que des composantes de la dynéine cytoplasmique et la dynactine, soit les gènes NudE, NudF et NudC (Morris, 2000). NudF est également connu sous le nom de LIS1. La protéine NudE et son isoforme NudEL se lient au niveau des chaînes intermédiaires et légères de la dynéine, et recrutent par la suite LIS1 en l'aidant à se positionner près du domaine moteur des hélices lourdes (Vallee *et al.*, 2012). La liaison de LIS1 permettrait une hausse de l'affinité de la dynéine pour les microtubules, permettant aux dynéines de rester stable sur les microtubules lors de transport de cargos de tailles substantielles (Vallee *et al.*, 2012).

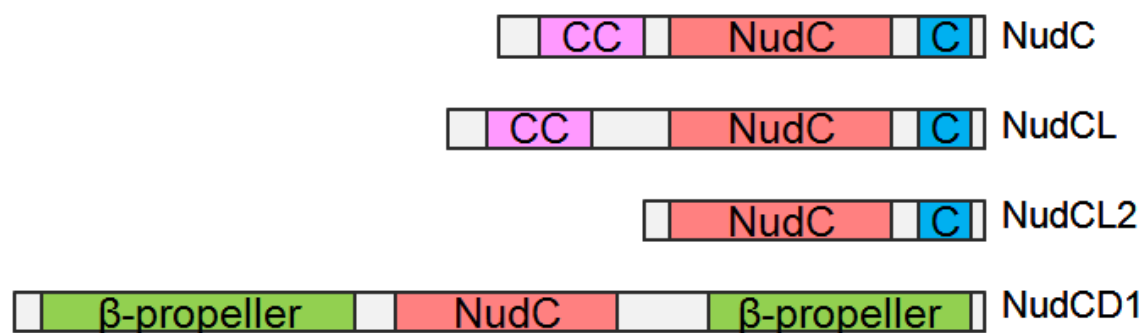


**Figure 7: Les protéines Nud et leurs interactions.** Schéma simplifié de l'interaction des protéines Nud faisant partie des complexes de la dynéine et dynactine. Les protéines NudE et NudF ne font pas partie de ces complexes, mais sont impliquées dans la régulation de ceux-ci.

### 1.3.2 NudC

La sous-famille de gènes NudC est hautement conservée chez différents organismes (Riera et Lazo, 2009). Cette sous-famille est constituée de 4 protéines : NudC, NudCL, NudCL2 et NudCD1 (Zheng *et al.*, 2011). Chaque membre de cette sous-famille possède un domaine ayant certaines séquences d'a.a. conservées, le domaine NudC/p23 (Figure 7). Ce domaine possède une structure suggérant une activité chaperone, agissant de manière indépendante ou de concert avec la chaperone HSP90 (Fu *et al.*, 2016). La plus étudiée, NudC, est impliquée dans différents processus cellulaires, tels que la mitose, le mouvement nucléaire, la réponse inflammatoire et la ciliogénèse (Zhang *et al.*, 2002, Aumais *et al.*, 2003, Riera et Lazo, 2009, Zhang *et al.*, 2016). NudC est également connue comme étant

impliquée dans certains cancers (Hartmann *et al.*, 2008). L'activité chaperonne de NudC lui permet entre autres de stabiliser la protéine LIS1, permettant la progression mitotique et ainsi la prolifération cellulaire (Zhu, 2010). NudCL, un paralogue de NudC, est nécessaire dans la progression du cycle cellulaire et la viabilité cellulaire en stabilisant la chaîne intermédiaire de dynéine (Zhou *et al.*, 2006). Un autre paralogue de NudC, NudCL2, stabilise la protéine LIS1 pour permettre le mouvement nucléaire (Yang *et al.*, 2010). Ces 3 membres possèdent une structure similaire, possédant tous le domaine NudC ainsi qu'un domaine commun C-terminal (Figure 8). NudC et NudCL possèdent également un motif superhélice en N-terminal, nécessaire pour la dimérisation de ces protéines (Figure 8) (Zheng *et al.*, 2011). La dimérisation NudC, NudCL et NudCL2 est indispensable au fonctionnement de l'activité hélicase orchestrée par le domaine NudC (Zheng *et al.*, 2011). NudCL2 ne possède pas ce domaine, mais parvient à homodimériser différemment grâce au dernier feuillet  $\beta$  de son domaine NudC et à son domaine en C-terminal (Figure 8) (Zheng *et al.*, 2011). Le quatrième membre de cette sous-famille, NudCD1, ne possède en commun que son domaine NudC, lequel est situé dans un motif  $\beta$ -propeller à 7 lames, constitué du reste de la protéine (Figure 8) (Zheng *et al.*, 2011).



**Figure 8: Structure moléculaire des protéines de la famille NudC.** Les protéines de la famille NudC possèdent toutes un domaine NudC/p23 en plus de domaines spécifiques à chaque protéine. CC : Superhélice. C: Domaine en hélices- $\alpha$

Les motifs  $\beta$ -propeller sont des assemblages de lames, celles-ci composées de 4 feuillets  $\beta$  antiparallèles, arrangés de manière à former un tunnel (Pons *et al.*, 2003). Ces motifs sont retrouvés dans un grand nombre de protéines, peuvent comporter de 3 à 10 lames et sont responsables de nombreuses fonctions biologiques différentes (Fülöp et Jones, 1999). Les



motifs à 7 lames sont retrouvés entre autres sur des protéines liant un ligand, des hydrolases, des lyases, des oxidoréductases, des protéines de signalisation, des protéines structurales et encore plus (Chen *et al.*, 2011). La fonction d'un domaine NudC inséré dans un motif à 7 lames n'est pas connue, mais pourrait lui donner une fonction chaperone spécifique (Faircloth *et al.*, 2008).

### 1.3.3 NudCD1

Contrairement aux autres membres de cette famille, NudCD1 ne semble pas être impliquée dans la migration nucléaire (Fu *et al.*, 2016). NudCD1, aussi connu sous le nom de CML66 ou OVA66, est une protéine de 66 kDa découverte dans le cadre d'une expérience visant à identifier des antigènes associés à la leucémie chez des patients atteints de leucémie myéloïde chronique (Wu *et al.*, 2000). Une expérience a détecté qu'au niveau des tissus sains, NudCD1 est exprimée à de très fort niveaux dans les testicules et à des niveaux plus faibles au niveau du coeur (Yang *et al.*, 2001). Cette même expérience a démontré la forte présence de NudCD1 dans différents types de cancers, autant dans des lignées cellulaires de tumeurs solides qu'hématopoïétiques (Yang *et al.*, 2001). Des anticorps contre NudCD1 ont été découverts dans 18 à 38% de patients atteints de cancers pulmonaires, prostatiques et de mélanomes, montrant son immunogénicité au niveau de différentes tumeurs solides en plus des leucémies (Yang *et al.*, 2001).

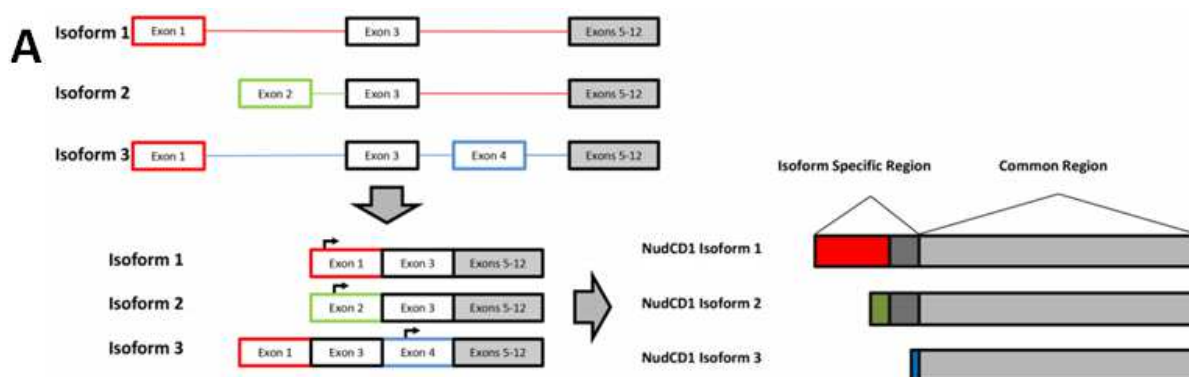
Une diminution de l'expression de NudCD1 sur des cellules HeLa, par ARN interférence, a démontré une diminution de la prolifération cellulaire *in vitro*, une atténuation des activités migratoires et invasives, ainsi qu'une augmentation de l'apoptose par rapport à des cellules non-transfectées (Wang *et al.*, 2008). Des expériences *in vivo* ont également démontré que la diminution de l'expression de NudCD1 inhibait la croissance tumorale et les métastases pulmonaires chez des souris nues inoculées avec des HeLa, par rapport aux souris contrôles ayant des HeLa où NudCD1 est exprimé normalement (Wang *et al.*, 2008). La sous-expression de NudCD1 mène à la diminution de plusieurs oncogènes : CTSL, MMP15, uPAR, VEGF, COX-2, S100A4, MUC1, MDM2 et RAC1 (Wang *et al.*, 2008). Une autre étude démontre que la surexpression de NudCD1 augmente la prolifération cellulaire sur des cellules non-cancéreuses ainsi que la résistance à l'apoptose induite par 5-FU (Rao *et al.*, 2014). Les auteurs de cette même étude ont effectué des xénogreffes de cellules NIH

3T3, des fibroblastes de souris, surexprimant ou n'exprimant pas NudCD1 chez des souris nues. Les xénogreffes de cellules surexprimant NudCD1 ont formé des tumeurs chez ces souris, en contraste avec les cellules n'exprimant pas NudCD1 qui n'ont pas formé de tumeurs, démontrant que l'expression de NudCD1 est oncogène *in vivo* (Rao *et al.*, 2014). La même étude a également montré que les phosphorylations des protéines AKT ou ERK1/2 sont nécessaires et que NudCD1 agissait probablement en amont de ces protéines de manière à favoriser leur phosphorylation (Rao *et al.*, 2014). Les mêmes auteurs ont démontré dans une autre étude que la diminution de l'expression de NudCD1 par ARN interférence sur des cellules de lignées tumorales menait à une réduction de l'activation de la voie de signalisation IGF1R-MAPK via l'expression de la protéine MDM2, proposant que NudCD1 agissait possiblement en amont de ce récepteur (Rao *et al.*, 2014). Ces mêmes cellules avaient également des niveaux réduits de phosphorylation sur les protéines Elk-1, Hsp27, Pyk2, ASK1, c-Kit, JNK, ATF2 et ERK1/2, qui sont connues pour être impliqués dans divers processus tumorigéniques, comme l'invasion, les métastases et l'apoptose (Rao *et al.*, 2014). Il a également été montré que la protéine peut interagir avec des composantes du complexe de transport vésiculaire COPI, ainsi qu'avec des membres de la famille des DEAD/DEAH box hélicases, impliqués dans différentes phases de la transcription et de l'épissage (Taipale *et al.*, 2014). NudCD1 possède également un potentiel immunothérapeutique, où des études démontrent la capacité de certains épitopes de la protéine à induire une réponse immunitaire dans différents types de leucémies (Suemori *et al.*, 2008, Suemori *et al.*, 2009)

#### 1.3.4 Isoformes de NudCD1

La majorité des connaissances sur NudCD1 sont au niveau de sa première isoforme. Le gène de NudCD1 est toutefois sous le contrôle de différents promoteurs et peut être épissé de manière à produire 3 différentes isoformes de la protéine (Yan *et al.*, 2004, Gerhard *et al.*, 2004). Le gène est constitué de 12 exons situés sur le chromosome 8 (Yang *et al.*, 2001, Yan *et al.*, 2004). Les exons 5 à 12 sont communs chez les 3 isoformes et elles possèdent chacune des combinaisons différentes des 4 premiers exons (Figure 9). La première isoforme de la protéine est constituée de 583 a.a. et est l'isoforme la plus étudiée et décrite. Les études et caractéristiques de NudCD1 précédemment démontrées font référence à cette

isoforme. La deuxième isoforme est constituée de 554 a.a., alors que la troisième en est composée de 496 (Yan *et al.*, 2004, Gerhard *et al.*, 2004). Les isoformes 1 et 2 ont en commun leurs 544 derniers a.a., alors que la troisième possède les mêmes 492 derniers a.a.. Ces isoformes se différencient donc par une région N-terminale spécifique (Figure 9). La deuxième isoforme est très fortement exprimée dans les tissus testiculaires sains et n'a pas été démontrée comme étant exprimée dans 4 différents types de lignées cellulaires tumorales testées (Yan *et al.*, 2004). Pour sa part, l'expression de l'isoforme 3 n'a pas été vérifiée, autant chez les tissus sains qu'au niveau de différents types de tumeurs.



**Figure 9: Isoformes de NudCD1.** Épissage alternatif du gène de NudCD1 menant à la formation d'isoformes de 3 tailles moléculaires différentes ayant une partie commune C-terminale et spécifique N-terminale.

#### 1.4 Problématique de recherche

Les études faites précédemment sur NudCD1 identifient principalement les effets de l'isoforme 1 de la protéine. Les effets et rôles que pourraient avoir les isoformes 2 et 3 ne sont donc pas connus, alors que leurs niveaux d'expression au niveau du cancer n'ont que peu été étudiés, seulement 4 lignées cellulaires pour l'isoforme 2, aucune pour l'isoforme 3 (Yan *et al.*, 2004). Également, malgré les connaissances sur les effets de surexpression et de sous-expression de NudCD1 dans la cellule, sa fonction et son mécanisme d'action sont toujours inconnus. En raison du manque de connaissances générales sur la deuxième et la troisième isoformes de NudCD1, ainsi que sur les fonctions que pourraient avoir les 3 différentes isoformes, cette étude **propose que les différentes isoformes de NudCD1**

**possèderaient des fonctions différentes dans le cancer.** De manière à valider cette hypothèse, trois objectifs principaux ont été fixés.

1. Mesurer l'expression des isoformes de NudCD1 dans différentes lignées cellulaires cancéreuses;
2. Déterminer la localisation des différentes isoformes de NudCD1 dans des lignées cellulaires;
3. Caractériser les partenaires d'interaction des isoformes de NudCD1.

Pour réaliser le premier objectif, l'expression des différentes isoformes a été mesurée par RT-PCR et immunobuvardage. Le second objectif a nécessité la génération de lignées cellulaires stables exprimant individuellement chaque isoforme avec des étiquettes. La localisation de NudCD1 dans ces lignées a par la suite été détectée par immunofluorescence. Le troisième objectif a nécessité l'utilisation des mêmes lignées cellulaires que pour le deuxième objectif. Des immunoprécipitations ou des purifications par affinité sur ces cellules ont été analysées en spectrométrie de masse afin de détecter un enrichissement protéique associé aux interactions avec les différentes isoformes de NudCD1.

## 2. MATÉRIEL ET MÉTHODES

### 2.1 Culture cellulaire

Les cellules U-2 OS, MIA PaCa-2, HEK 293T, HEK 293, HT-29, DLD-1, SW-1116, SW-480, SW-620, CaCo-2/15, HeLa et HCT 116 (Tableau 1) ont été cultivées dans du milieu d'Eagle modifié de Dulbecco (DMEM 1X, Life Technologies, Burlington, ON, Canada). Celui-ci a été complété avec 10% de sérum de veau fœtal (FBS) (Wisent, St-Bruno, QC, Canada), 100 U/mL de pénicilline et streptomycine (Life Technologies, Burlington, ON, Canada) et 2,5 µg/mL d'amphotéricin B (Life Technologies, Burlington, ON, Canada). Les cellules HIEC et CCD-18co (Tableau 1) ont été cultivées dans du milieu Opti-MEM (Life Technologies, Burlington, ON, Canada) supplémenté de 5% de FBS (Wisent, St-Bruno, QC, Canada), 100 U/mL de pénicilline et streptomycine (Invitrogen Life Technologies, Burlington, ON, Canada), 2,5 µg/mL d'amphotéricin B (Life Technologies, Burlington, ON, Canada), 1x d'acides aminés non-essentiels (Life Technologies, Burlington, ON, Canada), de 10 mM d'HEPES (Wisent, St-Bruno, QC, Canada) et 1x de GlutaMAX (Life Technologies, Burlington, ON, Canada). Les lignées stables ont été obtenues suite à l'intégration du vecteur pgLAP1, pgLAP5.2 ou pgLAP1-BirA (Tableau 2) dans des cellules U-2 OS Flp-In. L'induction des cellules a été faite avec 2 µg/mL de doxycycline (Clontech, Californie, USA) sur une période de 48 h. Les lignées cellulaires stables exprimant une protéine de fusion avec la biotine ligase sont également incubées pendant 24 h en présence de 50 mM biotine (Sigma-Aldrich Co, Oakville, ON, Canada) pour permettre la biotinylation des protéines.

Lignée cellulaire	Origine	Fournisseur
U-2 OS	Ostéosarcome humain	ATCC HTB-96
MIA PaCa-2	Carcinome humain	ATCC CRL-1420
HEK 293T	Rein embryonnaire humain transformé	ATCC CRL-11268
HEK 293	Rein embryonnaire humain	ATCC CRL-1573
HT-29	Adénocarcinome colorectal humain	ATCC HTB-38
DLD-1	Adénocarcinome colorectal humain de type Duke's C	ATCC CCL-221
SW-1116	Adénocarcinome colorectal humain de type Duke's A	ATCC CCL-233
SW-480	Adénocarcinome colorectal humain de type Duke's B	ATCC CCL-228
SW-640	Adénocarcinome colorectal humain de type Duke's C	ATCC CCL-227
CaCo-2/15	Adénocarcinome colorectal humain	ATCC HTB-37
HeLa	Adénocarcinome du col de l'utérus humain	ATCC CCL-2
HCT 116	Adénocarcinome colorectal humain	ATCC CCL-247
HIEC	Tissus normal intestinal de fœtus humain	ATCC CRL-3266
CCD-18co	Fibroblastes normaux du côlon humain	ATCC CRL-1459

**Tableau 1 : Lignées cellulaires utilisées.**

## **2.2 Lignées cellulaire stables, clonages et RT-PCR**

Les séquences des protéines NudCD1 isoformes 1, 2 et 3, ainsi que DHX15 ont été obtenues par amplification PCR d'ADNc obtenus par RT d'ARNs isolés de cellules U-2 OS (Annexe 2). Les RT ont été faites à partir d'un Petri de 35 mm et les ARNs ont été obtenus en suivant le protocole d'isolation au TRIzol® (Life Technologies, Burlington, ON, Canada). Pour l'obtention des ADNc, ceux-ci ont été obtenus en suivant le protocole de la ProtoScript® II *Reverse Transcriptase* (New England Biolabs, Massachusetts, USA). La polymérase utilisée pour l'obtention des ADNc est la iProof™ High-Fidelity (Bio-Rad

Laboratories, Ontario, Canada). Les conditions d'amplification ont été ainsi : une dénaturation initiale de 30 secondes à 98°C, une dénaturation de 10 secondes à 98°C, un appariement de 15 secondes à 68°C et une extension de 20 secondes à 72°C par 1000 paires de bases. Les trois dernières étapes étaient répétées 30 fois. Une extension finale était faite de 5 min à 72°C. Pour permettre l'utilisation de la technique de clonage utilisant des séquences recombinantes, des séquences attB ont été ajoutées aux protéines en N et en C-terminal. Les ADNc ont été insérés dans les vecteurs pDONR221 ou pDONR223 en utilisant le protocole de la BP Clonase II (Invitrogen Life Technologies, Burlington, ON, Canada). Par la suite, les séquences ont été transférées dans les différents vecteurs destinations (Tableau 2) à l'aide du protocole de la LR Clonase II (Invitrogen Life Technologies, Burlington, ON, Canada). Les RT-PCR déterminant l'expression protéique ont utilisé la même méthode, toutefois faite sur les différentes lignées cellulaires utilisées (Tableau 1) et les PCRs ont été migrées sur gel d'agarose 2% dans du tampon TAE 1x (40 mM Tris, 20 mM acide acétique, 1 mM EDTA à pH 8,3) pendant 40 min à 100V.

Plasmide	Étiquette	Type d'organisme	Fonction
pgLAP1	GFP	Eucaryote	Lignée stable
pgLAP5.2	GFP	Eucaryote	Lignée stable
pdLAP1-BirA	Myc et BirA	Eucaryote	Lignée stable
pDEST47	GFP	Eucaryote	Lignée transitoire
pDEST53-Myc	Myc	Eucaryote	Lignée transitoire

**Tableau 2 : Vecteurs d'expression utilisés**

### 2.3 Transfection transitoire et stable

Les transfections ont été faites en utilisant de la Lipofectamine LTX (Life Technologies, Burlington, ON, Canada). 5 µg d'ADN (ou 0,5 µg d'ADN et 4,5 µg du vecteur pOG44 codant pour la recombinaise dans le cas des transfections stables) et 5 µL de PLUS Reagent sont dilués dans 330 µL d'Opti-MEM avant que le mélange soit incubé 5 min. Un autre mélange de 5 µL de Lipofectamine LTX est incubé dans 330 µL d'Opti-MEM pour 5 min. Les deux mélanges sont combinés et incubés pour une période de 20 min à température ambiante. Le mélange est par la suite incorporé par goutelles sur des cellules à 50% de

confluence pour les transfections stables et 70% pour les transfections transitoires dans un Petri 60 mm. Les transfections transitoires sont lysées le lendemain alors que les lignées stables sont sélectionnées 48 h suivant la transfection. La sélection se fait avec 5 µg/mL de blasticidine (Wisent, St-Bruno, QC, Canada) et 10 µg/mL d'hygromycine (Wisent, St-Bruno, QC, Canada) pendant une semaine.

## **2.4 SDS-PAGE et immunobuvardage**

Les échantillons ont été resuspendus dans du tampon Laemli 5X (50% glycérol, 10% SDS, 300 mM Tris pH 6,8, bromophenol blue) avant d'être séparés par SDS-PAGE 10%. Les protéines sont ensuite transférées sur une membrane de nitrocellulose (PerkinElmer, Massachusetts, USA) et celle-ci est bloquée dans du TBS-T (10 mM Tris pH7,4, 150 mM NaCl, 0,1% Tween-20) avec 5% lait pendant 30 min à température ambiante. Les immunobuvardages utilisant la streptavidin-HRP pour leur part nécessitent un blocage dans une solution spécifique (PBS 1X, 0,4% triton X-100 et 2,5% BSA) pour une heure à température pièce. Les anticorps primaires (Tableau 3) sont dilués dans un tampon (1% BSA, 0,5% ovalbumin, 10 mM Tris pH7,4, 150 mM NaCl et 0,01% NaN<sub>3</sub>) et incubés pour 16 h à 4°C. Les membranes sont alors lavées 3 fois 10 min dans du TBS-T avant d'être incubées pour 2 h avec l'anticorps secondaire approprié (Tableau 3) dilué dans un tampon (2% BSA, 1% lait, 10 mM Tris pH7,4 et 150 mM NaCl) à température pièce. Les membranes sont ensuite lavées 3 fois 10 min avec le TBS-T avant d'être révélées avec du Clarity Western ECL (Bio-Rad Laboratories, Ontario, Canada) sur ChemiDoc Imaging System (Bio-Rad Laboratories, Ontario, Canada).



<b>Anticorps primaires</b>	<b>Dilution</b>	<b>Compagnie</b>	<b># de référence</b>
GFP	1/3000	Santa Cruz Biotech	Sc-8334
NudCD1	1/1000	Santa Cruz Biotech	Sc-138
GAPDH	1/50000	Cell Signalling	14C10
Streptavidin	1/20000	Life Technologies	19534-050
Myc/ Hybridome	1/50	ATCC	CRL-1729
BirA	1/1000	Sigma	GW20013F-50UG
SF3A3	1/1000	Abcam	56823
<b>Anticorps secondaires</b>	<b>Dilution</b>	<b>Compagnie</b>	<b># de référence</b>
Souris	1/10000	Santa Cruz Biotech	Sc-2005
Lapin	1/10000	Santa Cruz Biotech	Sc-2004
Poulet DyLight 650	1/500	Abcam	96954
Lapin Alexa Fluor 546	1/500	Invitrogen	A11010
Souris Alexa Fluor 488	1/500	Invitrogen	A11001

**Tableau 3: Anticorps primaire et secondaire utilisés.**

## **2.5 Immunofluorescence**

Les cellules U-2 OS ont été cultivées sur une lamelle avant d'être fixées avec du paraformaldéhyde 1% dans du PBS pour 10 min. Les cellules ont été ensuite perméabilisées avec du PBS contenant 0,5% Triton X-100 pour 10 min, les protéines sont marquées avec des anticorps primaires (Tableau 3) dilués 1/500 dans du PBS 1X. Suite à des lavages au PBS 0,1% Triton X-100 et ensuite au PBS, les cellules ont été marquées avec un anticorps secondaire couplé à l'Alexa 546, à l'Alexa 488 ou au DyLight 650 (Tableau 3). Après les lavages au PBS contenant 0,1% Triton X-100, les cellules ont été

colorées au DAPI 5 µg/ul (Sigma-Aldrich Co, Oakville, ON, Canada) avant lavages au PBS puis montage sur lames avec de liquide de montage et visualisation au microscope à 63x à l'huile.

## **2.6 Culture cellulaire par SILAC**

Les cellules U-2 OS Flp-In contrôle ont été cultivées dans du DMEM sans arginine et lysine contenant 5% de FBS triple-dialysé (Life Technologies Inc., Burlington, ON, Canada). L'arginine et la lysine ont été ajoutées au milieu comme isotopes légers (Arg0, Lys0; Sigma-Aldrich Canada, Oakville, ON, Canada). Les cellules U-2 OS exprimant les diverses constructions ont été cultivées dans le même DMEM contenant 5% de FBS triple-dialysé, mais avec des isotopes moyens (Arg6, Lys4; Cambridge Isotope Laboratories, Andover, MA, USA) et lourds (Arg10, Lys8; Cambridge Isotope Laboratories, Andover, MA, USA). La concentration finale de l'arginine est de 28 pg/mL et celle de la lysine est de 49 pg/mL. Ces cellules sont cultivées un minimum de 5 passages dans ces milieux afin de s'assurer une incorporation optimale des différents isotopes avant d'être amplifiées et traitées pour les expérimentations.

## **2.7 Immunoprécipitations et précipitations par affinité**

Lors de la méthode GFP, les 3 boîtes de Petri 150mm sont lavées 2 fois aux PBS 1X et les cellules sont lysées avec du tampon triton (1% Triton X-100, 50mM Tris pH 7,4, 150mM NaCl, 1% PMSF). Les extraits sont laissés sur glace pendant 10 min avant d'être soniqués à 3 reprises pour une période de 15 secondes à une intensité de 20% sur un sonicateur (Model 120 Sonic Dismembrator, Thermo Fisher Scientific Inc, Burlington, ON, Canada). Les échantillons sont alors centrifugés 10 min à 4000 g. Les billes GFP (GFP-Trap-A, ChromoTekGmbH, NY, USA), préalablement lavées 3 fois au tampon triton, sont ajoutées

En ce qui concerne la méthode BirA utilisant les billes de streptavidin, les 3 boîtes de Petri sont lavées 2 fois aux PBS 1X et les cellules sont lysées au tampon RIPA (25 mM Tris pH 7,4, 150 mM NaCl, 0,1% SDS, 0,5% sodium désoxycholate, 1% Triton X-100, 1% PMSF). Les extraits sont laissés sur glace pendant 10 min avant d'être soniqués à 3 reprises

pour une période de 15 secondes à une intensité de 20% sur un sonicateur (Model 120 Sonic Dismembrator, Thermo Fisher Scientific Inc, Burlington, ON, Canada). Les échantillons sont alors centrifugés 10 min à 4000 g. Les billes de Streptavidin High Performance (GE Healthcare, Buckinghamshire, UK), préalablement lavées 3 fois au tampon RIPA, sont ajoutées aux lysats et sont laissées 2 h sur rotateur à 4°C avant d'être centrifugées à 2000 g pendant 2 min et lavées 3 fois au tampon RIPA.

## **2.8 Digestions des protéines**

### **2.8.1 Digestion sur gel des protéines étiquetées GFP**

Les protéines immunoprécipitées sont réduites dans 10 mM de DTT, chauffées à 95°C 5 min puis alkylées dans 50 mM d'iodocétamide avant d'être séparées par SDS-PAGE (NuPAGE 4-12% Bis-Tris Gel, Novex, Thermo Fisher Scientific Inc, Burlington, ON, Canada). Le gel est coloré au SimplyBlue SafeStain (Invitrogen Life Technologies, Burlington, ON, Canada) pour visualiser les protéines. La ligne du puit est séparée en huit bandes et celles-ci sont coupées en plus petits morceaux avant d'être placées dans des tubes LoBind (Protein LoBind Tube 1,5 mL, Eppendorf, Mississauga, ON, Canada). Chaque bande est digérée avec de la trypsine (Trypsin Gold, Mass Spectrometry Grade, Promega Corporation, Winconsin, USA), diluée dans 20 mM  $\text{NH}_4\text{HCO}_3$ , à 30°C pour une période de 16 h. L'extraction des peptides se fait par l'ajout d'un volume égal de  $\text{CH}_3\text{CN}$  pour 30 min à 30°C avant que le surnageant soit transféré dans un nouveau tube Lobind. Les bandes sont réhydratées dans un volume d'acide formique 1% et incubées pendant 20 min à température ambiante. Le surnageant est transféré dans de nouveaux tubes et cette étape est répétée une autre fois. Les bandes découpées sont alors déshydratées dans du  $\text{CH}_3\text{CN}$  et les peptides du nouveau tube lyophilisés par SpeedVac. Ces peptides sont alors resuspendus dans 0,1% d'acide trifluoroacétique (TFA) (Sigma-Aldrich, St-Louis, MO, USA). Cette solution est alors passée sur ZipTip (EMD Millipore Corporation, ON, Canada) préalablement mouillé 3 fois avec 10  $\mu\text{L}$  de  $\text{CH}_3\text{CN}$  et ensuite équilibré 3 fois avec du TFA 0,1%. Les peptides sont passés sur ZipTip lors de 10 aspirations de haut en bas avec une pipette P10 avant 3 lavages de 10  $\mu\text{L}$  de TFA 0,1% et élution avec une solution 1% acide

formique et 50% CH<sub>3</sub>CN. Cette élution est alors lyophilisée et resuspendue dans 25 µL d'acide formique 1% (Fisher Scientific, Waltham, MA, USA).

### **2.8.2 Digestion sur billes des protéines étiquetées BirA**

Les protéines précipitées sont lavées 5 fois avec du NH<sub>4</sub>HCO 20 mM avant d'être réduites 30 min dans une solution de 10 mM DTT chauffé à 60°C et alkylées avec 15 mM iodoacétamide pendant 1 h à température ambiante. La concentration de DTT est ensuite augmentée à 15 mM puis les échantillons sont digérés à la trypsine pendant 16 h dans un tube LoBind. Cette réaction est arrêtée en ajoutant 1% d'acide formique et le surnageant est transféré dans un nouveau tube LoBind. Les billes sont resuspendues dans une solution de 60% CH<sub>3</sub>CN et 0,1% acide formique et centrifugées. Le surnageant est transféré dans le tube de l'étape précédente. Ces peptides sont lyophilisés au SpeedVac et resuspendus dans 0,1% TFA. Ces peptides sont alors resuspendus dans 0,1% d'acide trifluoroacétique (TFA) (Sigma-Aldrich, St-Louis, MO, USA). Cette solution est alors passée sur ZipTip (EMD Millipore Corporation, ON, Canada) préalablement mouillé 3 fois avec 10 µL de CH<sub>3</sub>CN et ensuite équilibré 3 fois avec du TFA 0,1%. Les peptides sont passés sur ZipTip lors de 10 aspirations de haut en bas avec une pipette P10 avant 3 lavages de 10 µL de TFA 0,1% et élution avec une solution 1% acide formique et 50% CH<sub>3</sub>CN. Cette élution est alors lyophilisée et resuspendue dans 25 µL d'acide formique 1% (Fisher Scientific, Waltham, MA, USA).

## **2.9 Chromatographie liquide couplée à la spectrométrie de masse en tandem (LC-MS/MS)**

Les peptides digérés ont été séparés par chromatographie en phase liquide à haute performance par le système Dionex Ultimate 3000 nanoHPLC. Un volume de 10 µL d'échantillon (environ 2 µg de peptides) dans 1% d'acide formique est injecté sur une colonne Acclaim PepMap100 C18 (0.3 mm id x 5 mm, Dionex Corporation, Thermo Scientific, Rockford, IL, USA) avec un flux de 4 µL/min. Suite à l'enrichissement, les peptides ont été élués sur une colonne PepMap C18 (75 µm x 50 cm, Dionex Corporation Thermo Scientific, Rockford, IL, USA) avec un gradient linéaire de 5-35% de solvant B

(90% acétonitrile avec 0,1% acide formique) sur une période de 240 min avec un flux de 200 nL/min. Le système de HPLC est associé à un spectromètre de masse Orbitrap QExactive (Thermo Fisher Scientific Inc, Burlington, ON, Canada) par une source EasySpray. Le voltage du vaporisateur a été ajusté à 2,0 kV et la température de la colonne a été ajustée à 40°C. L'acquisition des spectres en mode de balayage complet ( $m/z$  350-1600) a été enregistrée par l'Orbitrap avec une résolution de 70 000, après l'accumulation d'un milliard d'ions. Les dix ions peptidiques les plus intenses de l'analyse précédente de l'Orbitrap ont été fragmentés par dissociation induite de collision (l'énergie de collision normalisée à 35% et la résolution de 17 500) après l'accumulation de 50 000 ions. Le temps de remplissage était de 250 ms pour l'analyse complète et 60 ms pour l'analyse du MS/MS. Les charges des ions précurseurs ont été évaluées et tous les états de charge non-affectés ainsi que séparément, les ions chargés +7 ou +8, ont été rejetées. La liste d'exclusion dynamique a été restreinte à un maximum de 500 entrées avec un maximum de temps de rétention de 40 secondes et une possibilité de masse relative de 10 ppm. L'option de verrouillage des masses a été activée pour l'analyse afin d'améliorer la précision de la masse. Les données ont été acquises par le logiciel Xcalibur (Thermo Fisher Scientific, Waltham, É.-U.).

## **2.10 Quantification et analyse bio-informatique**

Les données ont été transformées, recherchées et quantifiées avec le logiciel MaxQuant version 1.4.1.2 avec l'utilisation de la banque de données Human Uniprot (16/07/2013, 88,354 entrées). Pour l'identification protéique, la carboamidomethylation des cystéines a été utilisée comme modification fixe alors que l'oxydation des méthionines et l'acétylation des protéines en N-terminal ont été utilisées comme modifications variables. Une valeur maximale de 2 mauvaises coupures a été acceptée. La tolérance de masse maximale a été fixée à 7 ppm pour l'ion précurseur et à 20 ppm pour les fragments d'ions. L'option de requantification (re-quantify) a été sélectionnée pour calculer le ratio isotopique lorsque les paires d'isotopes n'avaient pas été identifiées, ce qui est souvent observé dans des expériences d'immunoprécipitations à cause de l'absence de la protéine d'intérêt dans le contrôle. Afin d'obtenir une identification fiable, toutes les protéines ont été acceptées sur

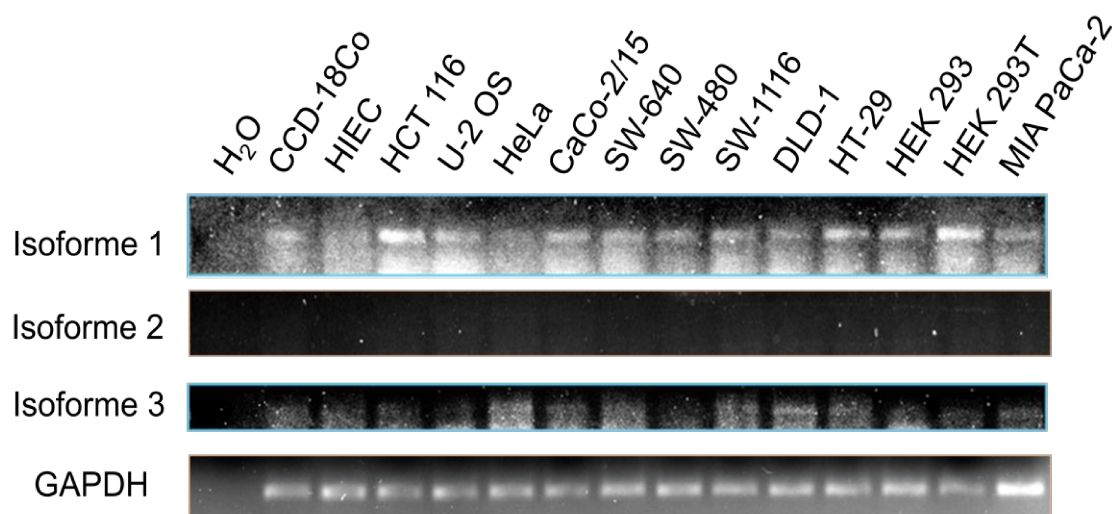
le critère que le nombre de protéines identifiées dans la base de données est 100 fois plus grand par rapport au nombre de protéines identifiées dans une base de données dont les séquences ont été inversées, menant ainsi à un taux de faux positifs de moins de 1 %. Pour la quantification de protéines, un minimum de deux peptides a été identifié pour chaque protéine. Les peptides provenant de différentes isoformes et les protéines qui ne peuvent pas être distinguées sur la base des peptides identifiés sont regroupés et affichés sur une seule ligne avec plusieurs numéros d'accession, comme c'est le cas pour NudCD1.

### 3. RÉSULTATS

#### 3.1 Niveaux d'expression des trois isoformes de NudCD1 dans différentes lignées cellulaires

##### 3.1.1 Par méthode de RT-PCR

Les études précédentes sur NudCD1 ont mesuré l'expression cellulaire de l'ARNm de la première isoforme dans différents types de cellules cancéreuses ainsi que la seconde isoforme dans quatre lignées cancéreuses (Yang *et al.*, 2001, Yan *et al.*, 2004). L'expression de l'isoforme 3 n'a, pour sa part, pas été mesurée. Des mesures des niveaux d'expression des différentes isoformes de NudCD1 dans différentes lignées cancéreuses humaines, particulièrement colorectales, ont été effectuées. Ainsi, des RT-PCR sur les lignées de cellules de cancer colorectales DLD-1, HT-29, HCT 116, CaCo-2/15, SW-620, SW-480, SW-1116, sur la lignée d'ostéosarcome U-2 OS, sur les lignées rénales embryonnaires HEK 293 et HEK 293T, sur la lignée de cancer cervical HeLa, sur la lignée de cancer pancréatique MIA PaCa-2 et finalement sur les lignées non-cancéreuses d'épithélium intestinal HIEC et de fibroblaste de côlon CCD-18co, ont été pratiqués. Les PCR ont été faites utilisant des amorces spécifiques aux isoformes 1 2 et 3 et ont été révélées sur gel d'agarose, elles démontrent une présence de modérée à forte de l'isoforme 1 dans toutes les lignées cancéreuses sauf les HeLa, et faible chez les lignées non-cancéreuses (Figure 10). La présence de l'isoforme 2 n'a pas été détectée et l'isoforme 3 est très faiblement présente dans toutes les lignées mesurées (Figure 10).

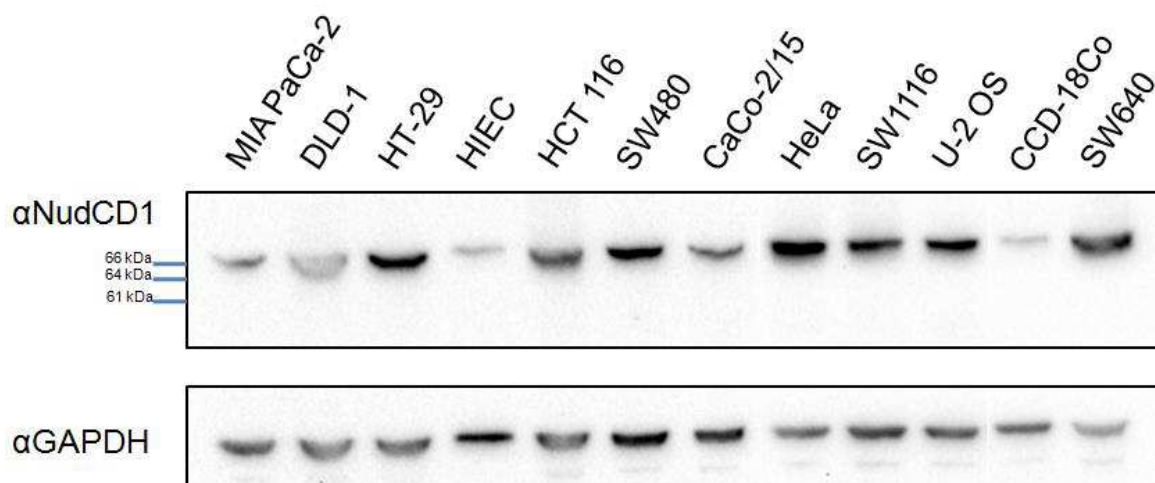


**Figure 10: Niveaux d'expression en ARNm des isoformes de NudCD1 dans différentes lignées cellulaires.** Expression d'ARNm cellulaire des trois isoformes de NudCD1 dans des lignées cellulaires cancéreuses et non-cancéreuses par RT-PCR semi-quantitatif migré dans un gel d'agarose 2%. Un contrôle négatif constitué d'eau a été utilisé lors des RT-PCR. Les transcrits de l'isoforme 1 et du contrôle positif GAPDH ont été examinés par RT-PCR de 30 cycles, alors que les transcrits de l'isoforme 2 et 3 ont été examinés sur une amplification de 35 cycles.

### 3.1.2 Par méthode d'immunobuvardage

Des études précédentes ont mesuré l'expression protéique de NudCD1 dans quelques lignées cellulaires cancéreuses non-colorectales (Rao, 2014) sans porter une attention particulière aux différentes isoformes. Les niveaux d'expression protéique des différentes isoformes de NudCD1 dans les lignées cellulaires énumérées précédemment, excluant les HEK 293 et HEK 293T, ont été mesurés. L'immunobuvardage sur les extraits protéiques des lignées cellulaires a démontré que toutes les lignées cancéreuses ont des niveaux d'expression protéique de l'isoforme 1 modérés à forts et que les lignées non-cancéreuses ont des niveaux d'expressions faibles de cette isoforme (Figure 11). Les isoformes 2 et 3 n'ont pas été détectées dans les lignées cancéreuses et non-cancéreuses malgré leur différence de poids moléculaire.



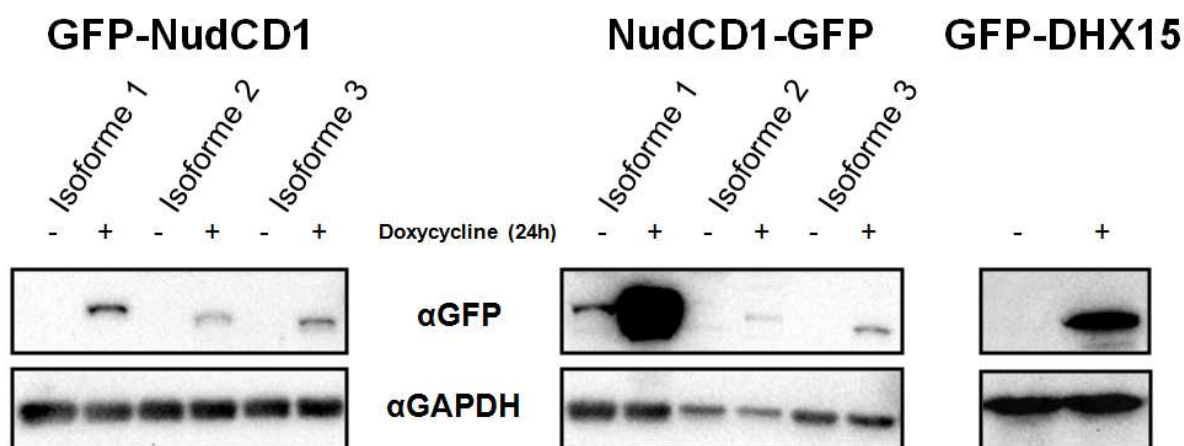


**Figure 11: Niveaux d'expression protéique des isoformes de NudCD1 dans différentes lignées cellulaires.** Expression protéique cellulaire des trois isoforme de NudCD1 dans des lignées cellulaires cancéreuses et non-cancéreuses par immunobuvardage avec anticorps contre NudCD1. La protéine GAPDH a été utilisée comme contrôle de chargement.

### 3.2. Génération de lignées stables de chaque isoforme de NudCD1

#### 3.2.1 Lignées stables avec étiquette GFP

Dans le but de pouvoir expérimenter sur chaque isoforme individuellement et sans contraintes de niveau d'expression ou de spécificité d'anticorps, des lignées de cellules U-2 OS exprimant des constructions de chaque isoforme de NudCD1 ayant des étiquettes GFP en C et en N terminale, ont été générées en utilisant les plasmides pgLAP1 et pgLAP5.2 respectivement. Le système Flp-In a été utilisé afin de n'avoir qu'un seul site d'intégration, et que l'expression de ces protéines soit inductible à la doxycycline. Les constructions cellulaires ont été validées par immunobuvardage à l'aide d'anticorps dirigés contre l'étiquette GFP des protéines. Les expérimentations démontrent l'expression attendue des isoformes de NudCD1 suite à induction à la doxycycline pour chaque isoforme et ce, pour les 2 types d'étiquettes (Figure 12). L'expression des isoformes 2 et 3 est plus faible, et la construction NudCD1-GFP pgLAP5.2 est exprimée très fortement, étant même exprimée en absence d'inducteur.

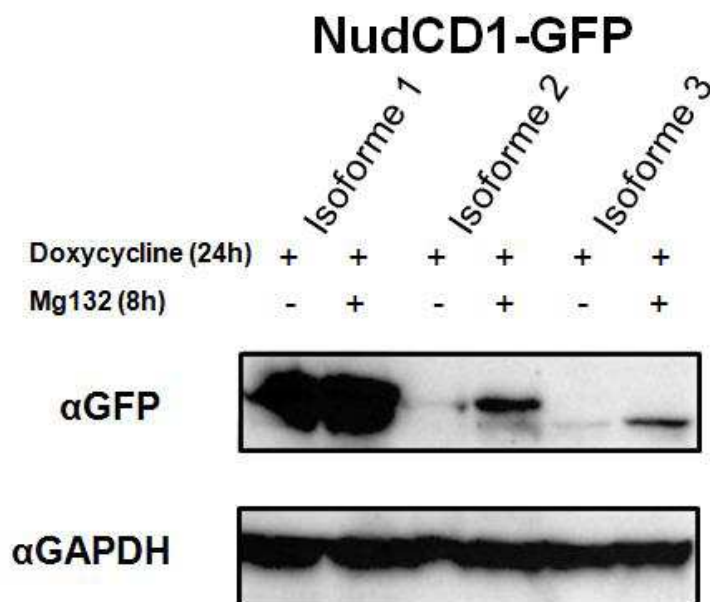


**Figure 12: Validation des constructions U-2 OS NudCD1 et DHX15 avec étiquette GFP.** Expression protéique des isoformes NudCD1 et de DHX15 étiquetées GFP par immunobuvardage. Chaque construction a été induite avec ou sans doxycycline pour 24h. La protéine GAPDH a été utilisée comme contrôle de chargement.

### 3.2.2 Lignées stables avec étiquette BirA

Des lignées cellulaires U-2 OS Flp-In exprimant la construction pgLAP5.2 NudCD1-BirA ont également été produites. Cette fois, les isoformes de NudCD1 ont été liées à une biotine ligase qui, lorsque mise dans un milieu contenant de la biotine, va produire des bioAMP qui vont interagir avec les lysines à la surface des protéines à proximité (Roux *et al.*, 2012). Les protéines ayant eu une interaction physiquement à proximité des protéines d'intérêt étiquetées seront donc biotinylées et pourront être purifiées par leur affinité avec la streptavidin, permettant des conditions de lyse dénaturantes et plus de lavages sans perdre d'interactions potentielles. Ces constructions ont été validées par immunobuvardage à l'aide de streptavidin-HRP, qui possède une forte affinité pour la biotine. L'isoforme 1 montre une forte biotinylation des protéines du lysat cellulaire, alors que les isoformes 2 et 3 montrent une très faible biotinylation des protéines par rapport aux contrôles sans doxycycline (Figure 13).



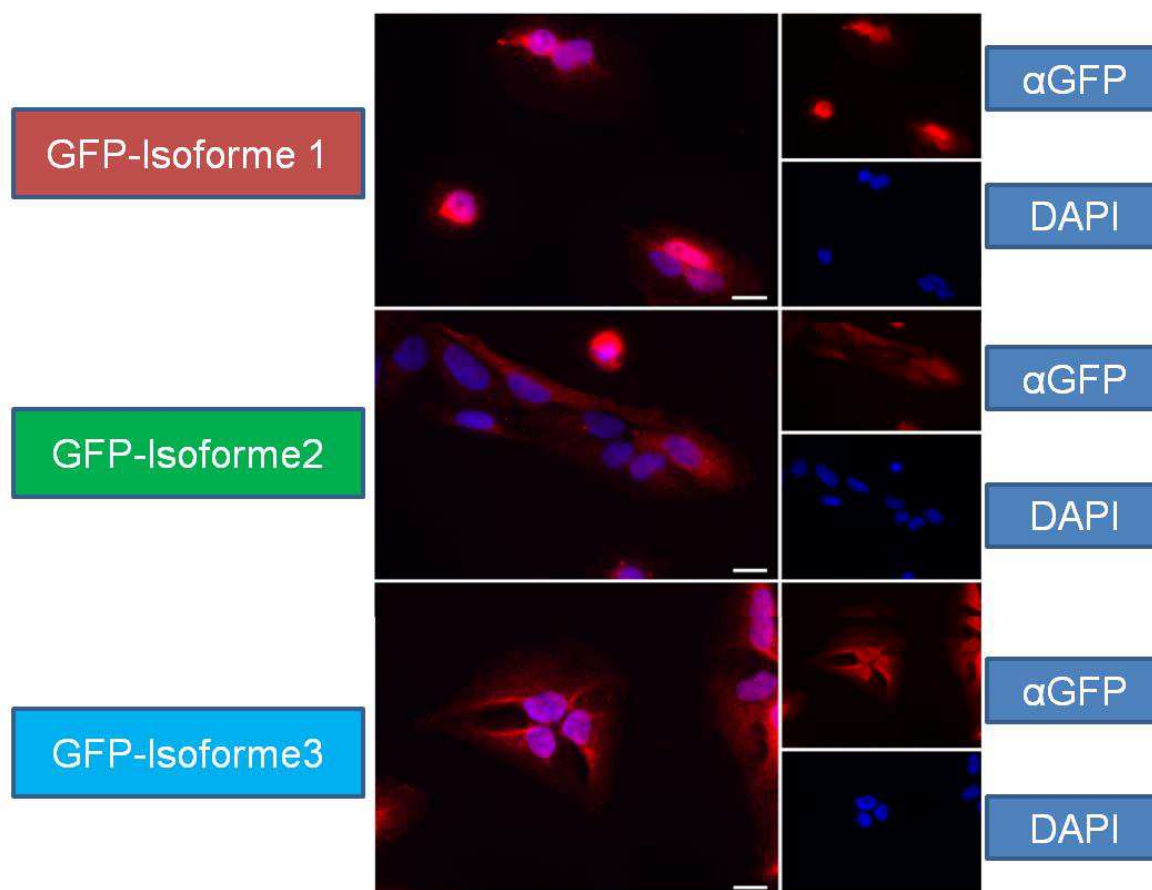


**Figure 14: Dégradation protéasomale des isoformes de NudCD1.** Les isoformes ont été traitées avec de la doxycycline pour 24 h avec ou sans MG132 pour une période de 8 h avant la lyse cellulaire. Les échantillons ont ensuite été passés par SDS-PAGE et immunobuvardage avant révélation avec un anticorps ciblant GFP. La protéine GAPDH a été utilisée comme contrôle de chargement.

### 3.4. Localisation des isoformes de NudCD1 dans la cellule

#### 3.4.1 Étiquette GFP N-terminale

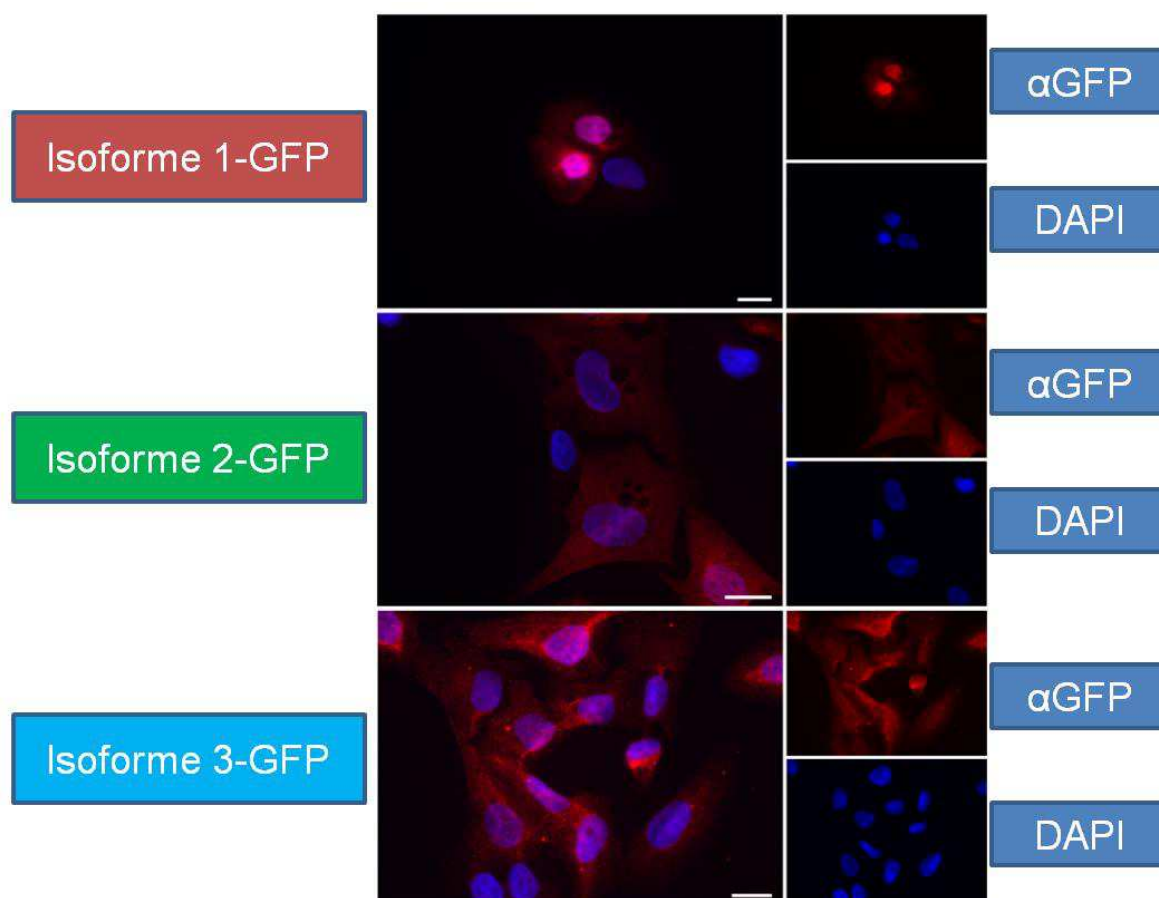
La localisation des différentes isoformes a été vérifiée par méthode d'immunofluorescence. Les constructions U-2 OS NudCD1 avec les étiquettes GFP en N et en C terminal ont été utilisées. Les immunofluorescences ont été effectuées en utilisant un anticorps contre GFP, permettant de cibler spécifiquement chaque isoforme et de distinguer si la présence en N ou en C-terminale de la construction GFP influence la localisation. Les immunofluorescences sur les cellules U-2 OS pgLAP1 ont montré que l'isoforme 1 est présente principalement au niveau du noyau des cellules, alors que les isoformes 2 et 3 semblent pour leur part être présentes au niveau du noyau et du cytoplasme des cellules (Figure 15).



**Figure 15: Localisation des isoformes de NudCD1 avec l'étiquette GFP en N-terminal.** Les cellules U-2 OS ayant les différentes constructions pgLAP1 NudCD1-GFP ont été traitées 24 h avec de la doxycycline avant préparation de lame. L'immunofluorescence a été effectuée avec un anticorps anti-GFP pour chaque isoforme et le DAPI a été utilisé afin de marquer les noyaux cellulaires. Les échelles représentent 10  $\mu$ m.

### 3.4.2 Étiquette GFP C-terminale

Les immunofluorescences sur les cellules U-2 OS pgLAP5.2, soit en C-terminal, ont également présentées une localisation nucléaire pour la première isoforme de NudCD1, une localisation majoritairement cytoplasmique pour NudCD1-2 ainsi qu'une localisation pancellulaire en ce qui concerne l'isoforme 3 (Figure 16). La présence en N et en C-terminal de l'étiquette GFP sur les différentes isoformes de NudCD1 ne modifie donc pas leur localisation cellulaire.

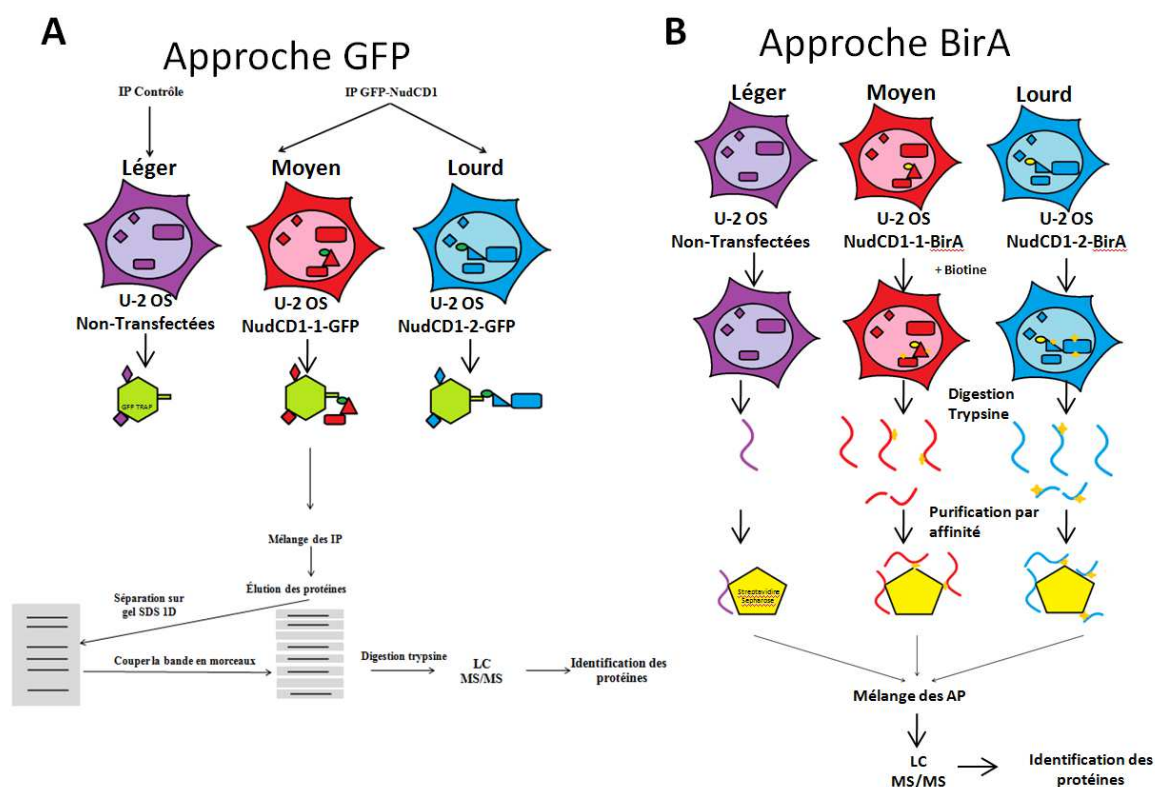


**Figure 16: Localisation des isoformes de NudCD1 avec l'étiquette GFP en C-terminal.** Les cellules U-2 OS ayant les différentes constructions pgLAP5.2 NudCD1-GFP ont été traitées 24 h avec de la doxycycline avant préparation de lame. L'immunofluorescence a été effectuée avec un anticorps anti-GFP pour chaque isoforme et le DAPI a été utilisé afin de marquer les noyaux cellulaires. Les échelles représentent 10  $\mu$ m.

### 3.5. Partenaires d'interaction de NudCD1

Dans l'optique de trouver les fonctions des isoformes de NudCD1, une approche de spectrométrie de masse en tandem couplée au SILAC, une méthode de protéomique quantitative, a été utilisée. Cette approche permet de comparer des niveaux de protéines immunoprécipitées par GFP-Trap dans un type de cellule, soit les cellules exprimant les différentes isoformes de NudCD1-GFP, avec ceux de cellules U-2 OS standard, n'exprimant pas de construction NudCD1-GFP, afin de discerner les partenaires d'interaction spécifiques des contaminants (Figure 17A). En plus de l'immunoprécipitation par GFP-Trap, la méthode de purification par billes de streptavidin a également été utilisée

pour les cellules exprimant les constructions NudCD1-BirA (Figure 17B). Ces expériences ont utilisé le milieu léger pour les U-2 OS n'exprimant pas les différentes constructions, comme contrôles négatifs, et les milieux moyens et lourds pour identifier les cellules exprimant les différentes isoformes de NudCD1. Des ratios moyen/léger ou lourd/léger ont été effectués pour déterminer l'enrichissement des protéines liant les isoformes par rapport aux contrôles. Seules les protéines ayant un ratio d'enrichissement supérieur ou égal à 2,0, et ce lors de deux expériences indépendantes, ont été considérées comme potentiels partenaires d'interaction.

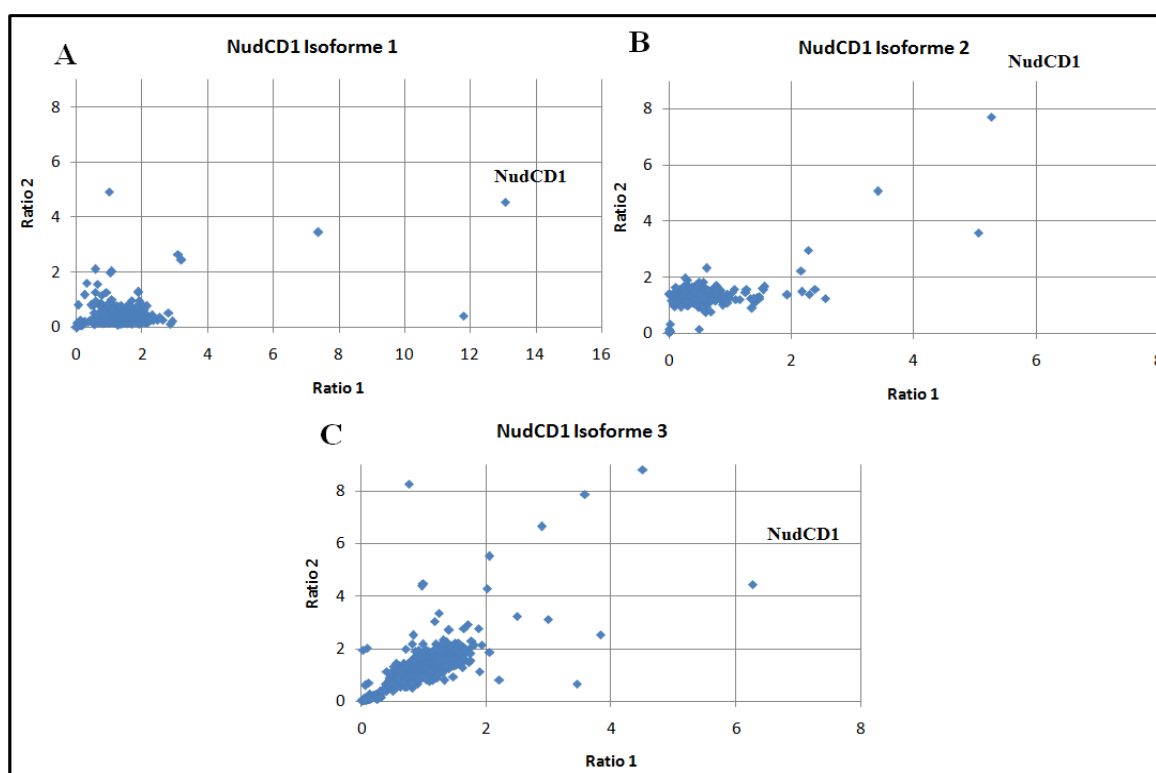


**Figure 17: Approches utilisées pour identifier des partenaires d'interaction aux isoformes de NudCD1.** Approche GFP en milieu SILAC avec les différentes isoformes de NudCD1 étiquetées GFP (A) et approche BirA en milieu SILAC où les différentes isoformes de NudCD1 étiquetées BirA ont interagi avec les traitements à la biotine (B).

Dans un premier cas, les immunoprécipitations par GFP-Trap des lysats de cellules exprimant les constructions des isoformes NudCD1-GFP ont été analysées. L'analyse sur l'isoforme 1 présente un total de 100 protéines ayant un ratio d'enrichissement d'au moins 2,0 réparties sur 2 expériences. Lorsque regroupées, il n'y a toutefois que 4 protéines s'y



retrouvant dans les 2 cas avec un ratio supérieur à 2,0, soit les protéines NudCD1, DHX15, COPB1 et GRWD1. DHX15 et GRWD1, en particulier, sont toutes deux impliquées dans l'épissage de l'ARN (Figure 18A, Tableau 4). L'analyse sur l'isoforme 2 présente pour sa part 28 protéines ayant un ratio suffisant, dont seulement 5 protéines qui s'y retrouvent à 2 occasions avec le ratio nécessaire, soit les protéines NudCD1, PSMD2, FECH, LOXL1 et SRCRB4D (Figure 18B, Tableau 4). Finalement, l'analyse sur la 3e isoforme donne 23 protéines, dont 9 respectent le biais, soit les protéines PSMD2, MED4, STK32B, NudCD1, EPPK1, MSFD9, SRCRB4D, PIK3C2B et LOXL1 (Figure 18C, Tableau 4). Un nombre de 3 protéines sont communes aux isoformes 2 et 3, soit les protéines PSMD2, LOXL1 et SRCRB4D. L'isoforme 1 ne possède aucune interaction commune avec les isoformes 2 et 3. Les protéines interagissant avec les isoformes 2 et 3 ne semblent pas avoir de voies ou de rôles direct associés, celles-ci étant impliquées dans diverses fonctions cellulaires, telles que la transcription et la réponse aux chocs thermiques.



**Figure 18: Partenaires d'interaction potentiels des isoformes de NudCD1 identifiés par spectrométrie de masse.** Ratios d'enrichissement des protéines immunoprécipitées sur des cellules exprimant l'isoforme 1(A), l'isoforme 2(B) et l'isoforme 3(C) étiquetées GFP par rapport au contrôle.



Isoforme 1		Isoforme 2		Isoforme 3			
Protéine	Ratio	Protéine	Ratio	Protéine	Ratio	Protéine	Ratio
NUDCD1	8,81	NUDCD1	5,27	PSMD2	6,66	MFSD9	3,18
DHX15	5,41	PSMD2	5,06	MED4	5,72	SRCRB4D	3,15
GRWD1	2,83	FECH	3,41	STK32B	5,35	PIK3C2B	3,05
COPB1	2,89	LOXL1	2,28	NUDCD1	4,77	LOXL1	2,87
		SRCRB4D	2,16	EPPK1	3,78		

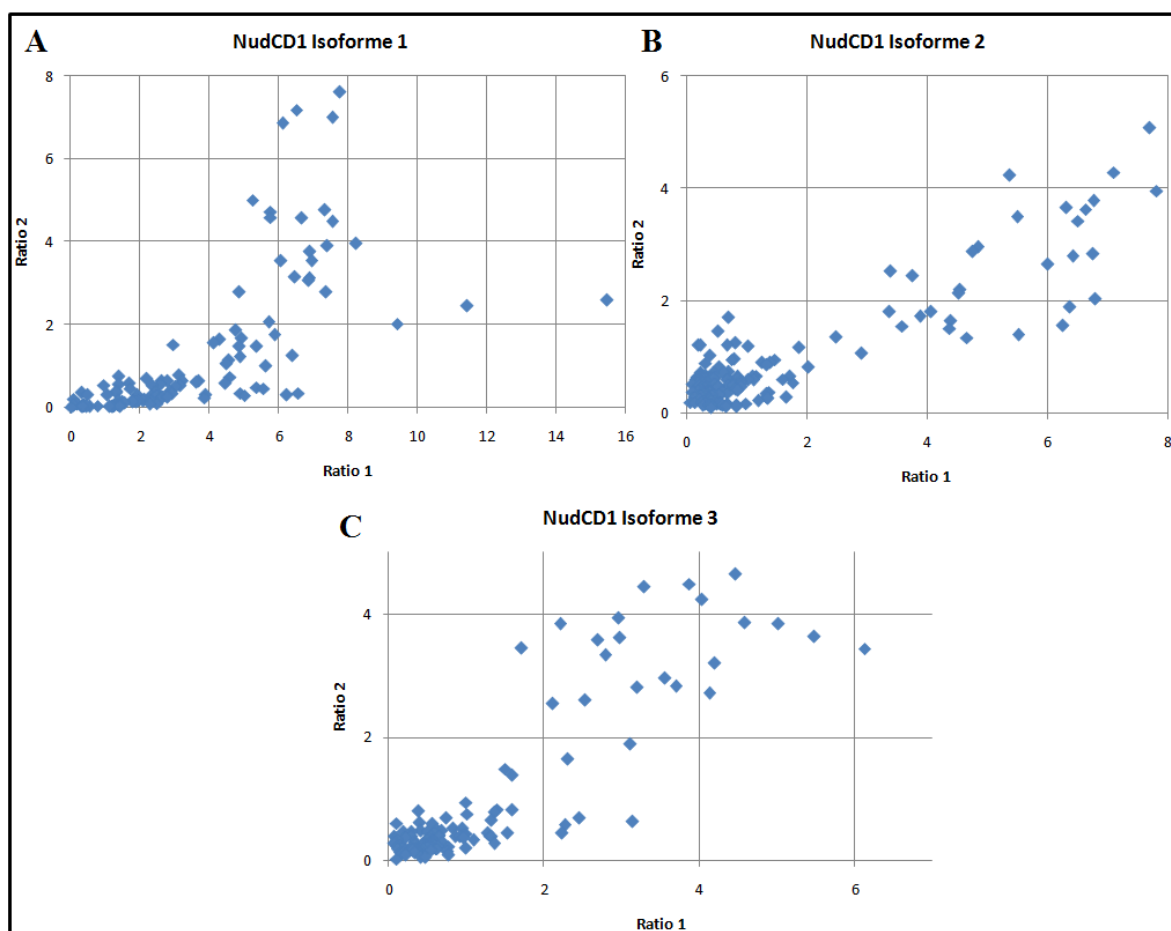
  

Isoforme 1		Isoforme 2		Isoforme 3			
Protéine	Ratio	Protéine	Ratio	Protéine	Ratio	Protéine	Ratio
NUDCD1	8,81	NUDCD1	5,27	PSMD2	6,66	MFSD9	3,18
DHX15	5,41	PSMD2	5,06	MED4	5,72	SRCRB4D	3,15
GRWD1	2,83	FECH	3,41	STK32B	5,35	PIK3C2B	3,05
COPB1	2,89	LOXL1	2,28	NUDCD1	4,77	LOXL1	2,87
		SRCRB4D	2,16	EPPK1	3,78		

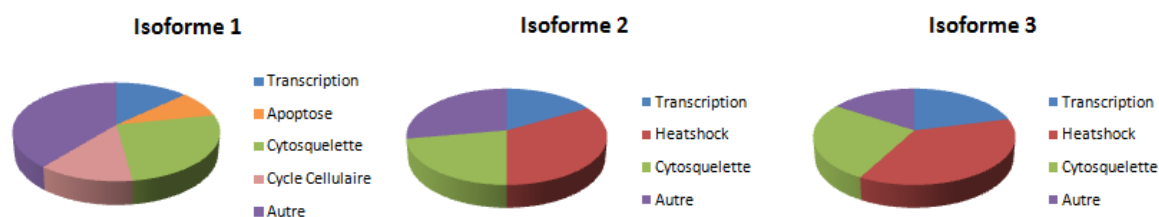
**Tableau 4 : Liste des protéines interagissant avec chacune des isoformes de NudCD1-GFP analysées au spectromètre de masse avec un ratio supérieur à 2,0.**

Dans un deuxième cas, les purifications par billes de streptavidin des lysats de cellules exprimant les constructions des isoformes de NudCD1-BirA ont été analysées. Les analyses ont montré que 24 protéines possédaient un ratio supérieur à 2,0 pour l'isoforme 1, 19 protéines pour l'isoforme 2 et 20 protéines pour l'isoforme 3 (Figure 19). Parmi celles-ci, une protéine est commune aux isoformes 1 et 2 : EPB41L3 et une autre protéine est partagée entre l'isoforme 1 et 3 : PDLIM4 (Tableau 5). Les isoformes 2 et 3 ont en commun 7 protéines : NUDC, MKL2, RUVBL1, RUBVL2, STIP1, HSP90AB1 et DNAJA1 (Tableau 5). Les trois isoformes possèdent 7 protéines interagissant en commun, soit les protéines MAP4, EEF1D, FLNA, TAGLN2, FASN, CCT8 et CORO1B (Tableau 5). L'isoforme 1 interagit principalement avec des protéines impliquées au niveau du cytosquelette et de l'apoptose, le reste étant des protéines individuellement impliquées dans d'autres voies, telles que l'épissage de préARNm, le transport vésiculaire ou la régulation du cycle cellulaire (Figure 20). Les isoformes 2 et 3, pour leur part, interagissent majoritairement avec des protéines impliquées au niveau de la régulation en réponse aux chocs thermiques, du cytosquelette ou de la transcription (Figure 20). Les deux méthodes

n'ont permis l'identification commune que d'une seule protéine, DHX15, interagissant avec l'isoforme 1 de NudCD1.



**Figure 19: Partenaires d'interaction potentiels des isoformes de NudCD1 identifiés par spectrométrie de masse.** Ratios d'enrichissement des protéines purifiées par affinité sur des cellules exprimant l'isoforme 1(A), l'isoforme 2 (B) et l'isoforme 3 (C) étiquetées BirA par rapport au contrôle.



**Figure 20: Regroupement fonctionnel des protéines ayant interagi avec les isoformes de NudCD1-BirA analysées au spectromètre de masse.** Protéines possédant un ratio d'enrichissement supérieur à 2,0.

Isoforme 1							
Protéine	Ratio	Protéine	Ratio	Protéine	Ratio	Protéine	Ratio
PCNP	9,04	TIPRL	6,10	CD2AP	5,33	CCT8	5,01
CTTN	7,70	HCFC1	6,05	AHNAK2	5,25	AHNAK	4,98
COPG2	7,27	EPB41L3	6,02	FLNA	5,24	PAICS	4,81
EFHD2	6,94	PDLIM4	5,72	TAGLN2	5,18	CORO1B	4,79
NUDCD1	6,85	MAP4	5,65	EIF4G1	5,13	IFIT5	3,88
DHX15	6,49	EEF1D	5,62	FASN	5,08	CTPS1	3,83

Isoforme 2			
Protéine	Ratio	Protéine	Ratio
DNAJA2	6,39	CORO1B	4,50
NUDC	5,88	FASN	4,41
MKL2	5,69	EPB41L3	4,32
RUVBL1	5,27	DNAJA1	3,90
FLNA	5,13	NUDCD1	3,82
RUVBL2	4,98	TAGLN2	3,37
STIP1	4,96	MAP4	3,33
CCT8	4,81	EIF2S1	3,10
HSP90AB1	4,79	HSPH1	2,96
EEF1D	4,60		

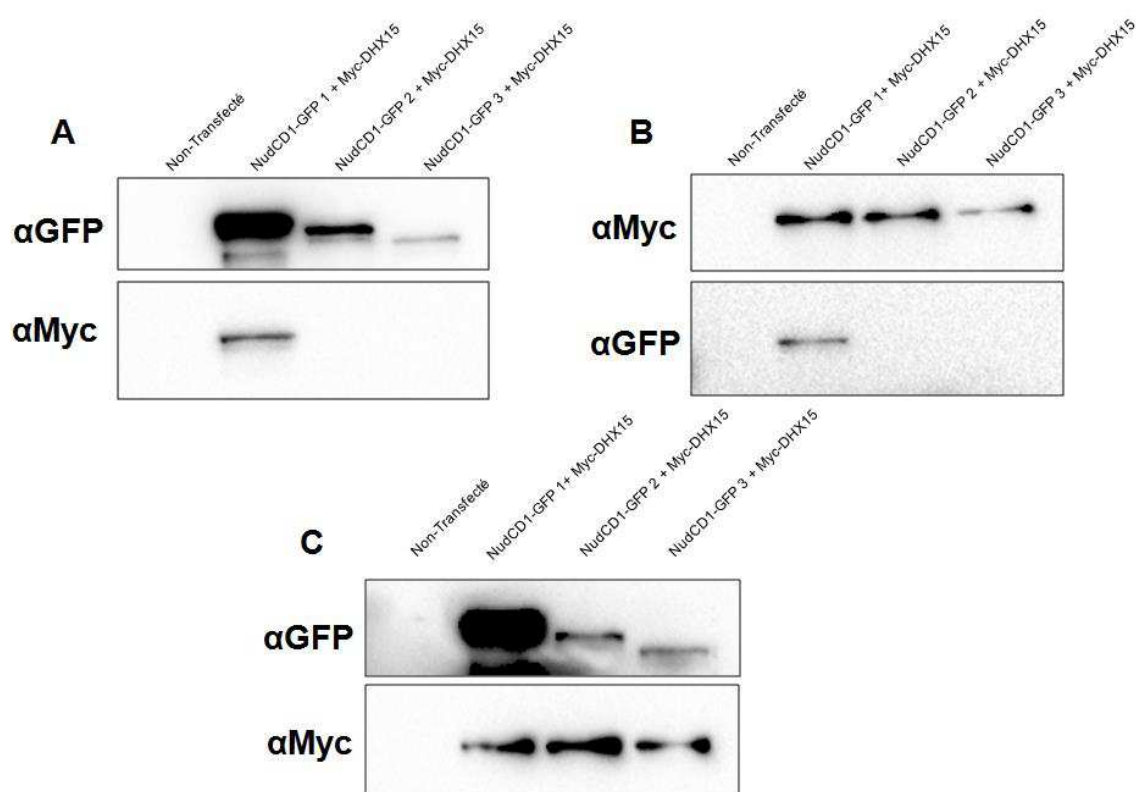
Isoforme 3			
Protéine	Ratio	Protéine	Ratio
FKBP4	4,78	HSPA1A	3,43
RUVBL1	4,56	CORO1B	3,30
NUDC	4,55	TAGLN2	3,26
HSP90AB1	4,42	DNAJA1	3,26
EEF1D	4,23	NUDCD1	3,13
RUVBL2	4,18	FASN	3,06
CCT8	4,14	MKL2	3,02
FLNA	3,86	PDLIM4	3,01
STIP1	3,69	EEF1G	2,56
MAP4	3,44	HSPA8	2,33

**Tableau 5 : Liste des protéines interagissant avec chacune des isoformes de NudCD1-BirA analysées au spectromètre de masse avec un ratio supérieur à 2,0.**

### 3.6. La protéine DHX15 interagit avec l'isoforme 1 de NudCD1

Pour confirmer l'interaction NudCD1 isoforme 1 avec DHX15, la seule protéine interagissant significativement trouvée en utilisant les deux méthodes de spectromètre de masse, des cellules HEK 293T ont été co-transfectées avec les isoformes 1, 2 ou 3 de NudCD1-GFP et des constructions Myc-DHX15. Dans un premier cas, une immunoprécipitation contre l'étiquette GFP a été effectuée. L'immunoprécipitation a été réduite et alkylée avant d'être migrée par SDS-PAGE. Un immunobuvardage de cette immunoprécipitation avec des anticorps contre les étiquettes Myc et GFP pour révéler les interactions (Figure 21A). La co-transfection de l'isoforme 1 et de DHX15 est la seule ayant révélé autant le NudCD1-GFP que le Myc-DHX15. À l'opposé, nous n'avons pas observé d'interaction entre les isoformes 2 et 3 et la protéine DHX15. Parallèlement, une immunoprécipitation contre l'étiquette Myc a été effectuée. Cette immunoprécipitation a été révélée avec les anticorps contre Myc et GFP. La co-transfection NudCD1 isoforme 1

et DHX15 est la seule qui a révélé une bande au niveau de l'anticorps GFP (Figure 21B). Les deux immunoprécipitations ont montré la co-immunoprécipitation de l'isoforme 1 et DHX15, alors que les isoformes 2 et 3 n'ont pas montré d'interaction avec DHX15.

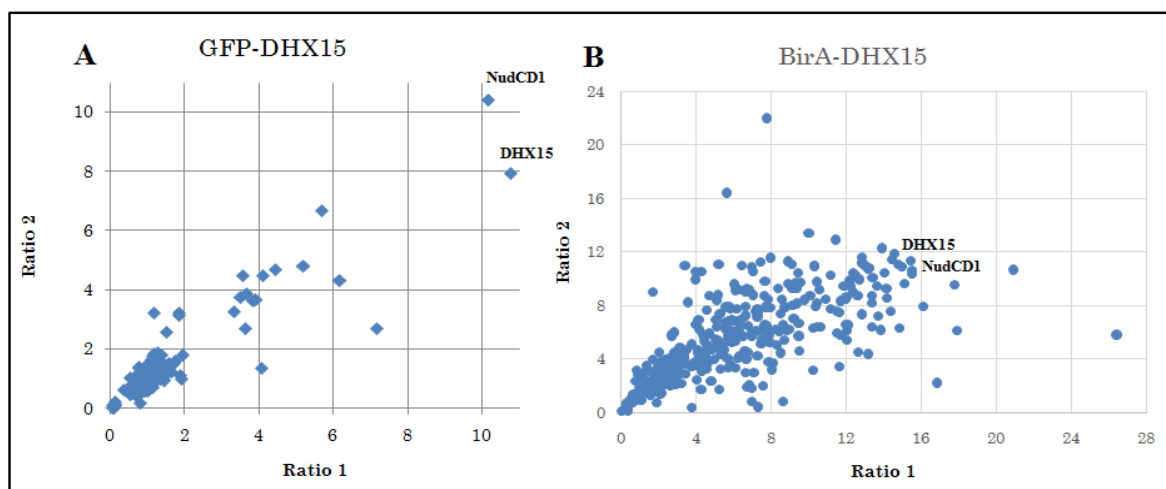


**Figure 21: Interaction entre l'isoforme 1 de la protéine NudCD1 et la protéine DHX15.** Des extraits de lignées cellulaires HEK293T ont été co-transfectés avec des constructions NudCD1-GFP isoformes 1 à 3 et une construction Myc-DHX15. Immunoprécipitation de GFP et révélation contre GFP et Myc(A). Immunoprécipitation de Myc et révélation contre Myc et GFP(B). Extraits totaux et révélation contre Myc et GFP (C).

### 3.7. Partenaires d'interaction de DHX15

Afin de confirmer l'interaction NudCD1-DHX15 et de confirmer la fonction de DHX15, la méthode SILAC couplée au spectromètre de masse a encore une fois été utilisée. Les conditions expérimentales ont cette fois-ci impliqué des cellules U-2 OS GFP-DHX15 ou U-2 OS BirA-DHX15 dans le milieu lourd, en comparaison avec le contrôle U-2 OS en milieu léger (Figure 22). Les protéines ayant un ratio d'enrichissement supérieur ou égal à

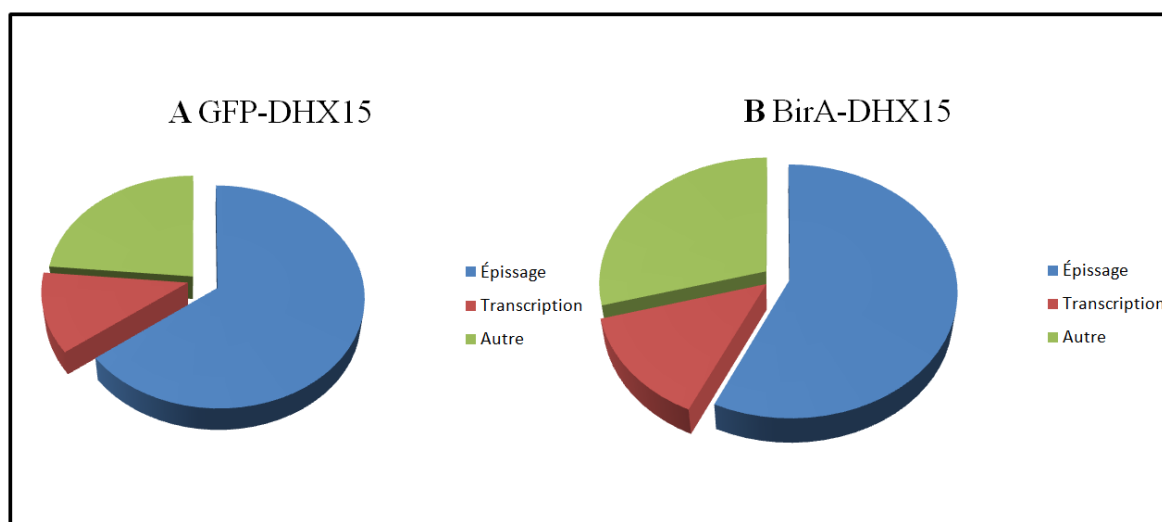
2,0 pour l'expérience GFP ont été considérées comme potentiels interacteurs. Dans le cas de l'expérience utilisant l'étiquette BirA, en raison d'un très grand nombre de protéines identifiées, seules les protéines ayant un ratio supérieur à 8,0 ont été considérées. L'expérience sur GFP-DHX15 a identifié 16 protéines interagissant avec DHX15 avec un niveau d'enrichissement nécessaire (Tableau 6). Parmi ces protéines, 10 sont impliquées dans des mécanismes d'épissage d'ARNm et 4 d'entre elles sont impliquées dans la transcription (Figure 23). On retrouve dans la première catégorie, des protéines telles que SF3A1, SF3B1 et TFIP11 qui sont des membres connus du spliceosome. On retrouve la protéine NudCD1 présente avec le plus haut ratio d'enrichissement autre que la protéine DHX15. Dans l'expérience sur BirA-DHX15, 58 protéines ont un ratio supérieur à 8,0. Un nombre de 33 de ces protéines sont impliquées dans l'épissage d'ARNm, 8 d'entre elles dans la transcription et le reste ont des fonctions autres ou inconnues (Figure 23). Parmi les 58 protéines, NudCD1 est la 9e la plus enrichie et donne une autre preuve de l'interaction.



**Figure 22: Partenaires d'interaction potentiels de DHX 15 identifiés par spectrométrie de masse.** Ratios d'enrichissement des protéines immunoprécipitées sur des cellules exprimant GFP-DHX15(A) et BirA-DHX15(B) par rapport au contrôle.

GFP-DHX15		BirA-DHX15					
Protéine	Ratio	Protéine	Ratio	Protéine	Ratio	Protéine	Ratio
<b>DHX15</b>	10,285	GPATCH8	18,801	<b>TFIP11</b>	11,636	SNRPN	10,465
<b>NUDCD1</b>	9,3675	FNBP4	14,051	<b>CHERP</b>	11,583	RBM26	10,418
<b>PAXBP1</b>	6,1847	CPSF6	14,016	RBM10	11,560	SART1	10,385
GCFC2	5,2514	<b>U2SURP</b>	13,606	<b>SF3B1</b>	11,506	IK	10,235
<b>TFIP11</b>	5,0075	<b>SF3B2</b>	13,513	PHAX	11,435	GTF2E1	10,134
CMTR1	4,9314	DDX46	13,073	NASP	11,407	NAT10	9,9020
SF3A3	4,5655	PSME3	12,955	HCFC1	11,294	EXOSC10	9,5532
<b>SF3A1</b>	4,2981	XRN2	12,930	RBM27	11,258	TRPC3	9,5463
SNRPA1	4,0182	<b>NUDCD1</b>	12,642	<b>RBM17</b>	11,241	RUVBL2	9,5143
<b>SF3B2</b>	3,8014	DDX42	12,319	SUGP2	11,192	PRPF3	9,3928
<b>CHERP</b>	3,7682	<b>PAXBP1</b>	12,245	TOX4	11,090	PRPF19	9,3920
<b>U2SURP</b>	3,7340	<b>SF3B3</b>	12,223	THRAP3	10,936	WAPAL	9,3181
<b>RBM17</b>	3,6385	EXOSC9	12,207	ANLN	10,916	FIP1L1	9,1855
<b>SF3B1</b>	3,2988	ZBTB33	12,087	<b>SF3A1</b>	10,906	FUBP1	9,0855
<b>SF3B3</b>	3,1497	NPM1	11,965	CPSF1	10,899	SAP30BP	8,8418
POLDIP2	2,5406	NUP50	11,961	TCERG1	10,791	KPNA1	8,5910
LANCL1	2,4989	CCDC174	11,875	TAGLN2	10,761	RBM25	8,5110
		<b>DHX15</b>	11,765	SON	10,671	RBBP4	8,5102
		NUDT21	11,715	RBM5	10,653	SNRPD1	8,4537

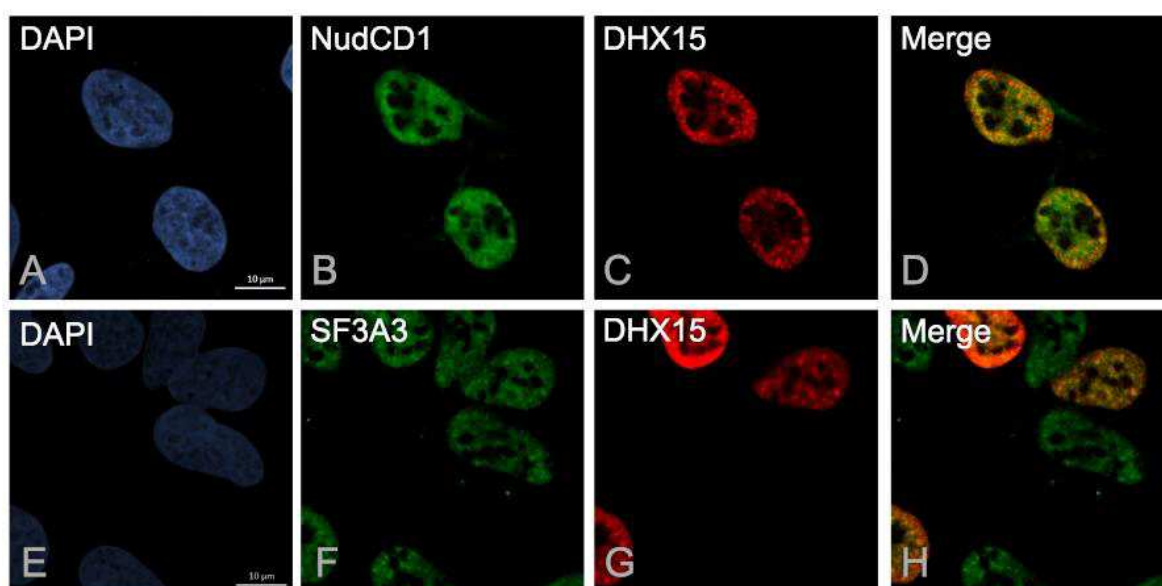
**Tableau 6 : Liste des protéines interagissant avec DHX15 identifiées par spectrométrie de masse.** Protéines analysées au spectromètre de masse avec un ratio supérieur à 2,0 pour GFP et 8,0 pour BirA. Les protéines en gras ont été identifiées lors des 2 expériences.



**Figure 23: Regroupement fonctionnel des protéines ayant interagi avec DHX15 identifiées par spectrométrie de masse.** Protéines analysées au spectromètre de masse avec un ratio supérieur à 2,0 pour GFP(A) et 8,0 pour BirA(B).

### 3.8. NudCD1 colocalise avec DHX15

La localisation de DHX15 a été vérifiée par immunofluorescence. Des cellules U-2 OS ont été co-transfectées avec des constructions BirA-DHX15 et de l'isoforme 1 de NudCD1-GFP avant d'être préparées et révélées à l'aide d'un anticorps anti-GFP et anti-BirA pour permettre la détection de NudCD1 et DHX15 respectivement. DHX15 et l'isoforme 1 ont une localisation nucléaire. En fusionnant les images, on détecte une co-localisation de NudCD1 et DHX15 à certain endroits. Des immunofluorescences ont également été faites pour vérifier la colocalisation de DHX15 et SF3A3, un facteur d'épissage nécessaire pour l'assemblage du pré-spliceosome, détecté précédemment dans les expériences de spectrométrie de masse. Des cellules U-2 OS ont été transfectées avec BirA-DHX15 puis préparation de cellules sur lame. Les révélations des protéines BirA-DHX15, via un anticorps anti-BirA, et de SF3A3, à l'aide d'un anticorps ciblant directement la protéine, ont démontré également une co-localisation de ces deux protéines.



**Figure 24: Co-localisation de DHX15 avec NudCD1 ou le facteur d'épissage SF3A3 par immunofluorescence dans des cellules U-2 OS.** Les noyaux ont été détectés par DAPI (A et E). Certain points de co-localisation entre NudCD1(B) et DHX15(C), ainsi que SF3A3(F) et DHX15 (G) ont été détectés dans le noyau (D et H, respectivement).

## 4. DISCUSSION

L'expression du gène de NudCD1 a initialement été observée chez des patients atteints de leucémie myéloïde chronique comme étant un nouveau gène identifié et caractérisé (Wu *et al.*, 2000). Ce gène peut être épissé alternativement de façon à produire 3 différentes isoformes ayant une région C-terminale commune et une région N-terminale spécifique (Yan *et al.*, 2004, Gerhard *et al.*, 2004). Des études de caractérisation sur la première isoforme ont ainsi démontré que la protéine était impliquée dans différents types de tumeurs en agissant possiblement de manière à activer IGF1R et ainsi favoriser la tumorigénèse et survie cellulaire par les voies de signalisation d'AKT et des MAPK et qu'elle n'était exprimée normalement qu'au niveau des testicules et plus faiblement, au coeur (Yang *et al.*, 2001, Rao *et al.*, 2014). Les deux autres isoformes n'ont pour leur part pas été caractérisées. L'absence de connaissances sur ces isoformes, ainsi que les effets de l'expression de la première isoforme qui mènent à l'apparition de phénotypes tumoraux, a mené à la proposition que les différentes isoformes de NudCD1 auraient des fonctions différentes dans le cancer.

### 4.1 Les isoformes de NudCD1 ne sont pas toutes exprimées dans les différentes lignées cellulaires cancéreuses.

Peu de recherches ont été faites sur le statut d'expression des isoformes 2 et 3 dans le cancer. L'expression de l'isoforme 3 n'a pas été vérifiée alors que l'expression de l'isoforme 2 a été confirmée dans des tissus testiculaires et démontrée comme étant très faible chez 4 lignées cellulaires cancéreuses, toutefois aucune d'entre elles ne provenant de cancer colorectal (Yan *et al.*, 2004). L'un des objectifs était donc de mesurer l'expression des trois isoformes dans différentes lignées pour confirmer la littérature, ainsi que de déterminer le niveau d'expression de la 3e isoforme dans différentes lignées cellulaires. Des expériences de RT-PCR et d'immunobuvardage de type western ont été effectuées afin de mesurer les niveaux d'expression des différentes isoformes de NudCD1 chez différentes lignées cellulaires cancéreuses et non-cancéreuses. Les RT-PCR montrant l'expression d'ARNm ont démontré la présence de l'isoforme 1 à différents niveaux d'expression chez toutes les lignées cellulaires, incluant les HIEC et les CCD-18co, deux lignées non-cancéreuses (Figure 9). Les niveaux d'expression sont toutefois plus faibles chez ces deux lignées



cellulaires, en plus de la lignée tumorale HeLa. Pour sa part, aucune expression d'ARNm de l'isoforme 2 n'a été mesurée chez ces mêmes cellules. Des transcrits de l'isoforme 3 ont été mesurés dans la majorité des lignées cellulaires, toutefois dans des niveaux d'expression inférieurs à l'isoforme 1, ce malgré 35 cycles de polymérisation pour les isoformes 2 et 3 par rapport à 30 cycles pour l'isoforme 1. Pour confirmer les différents niveaux d'expression des isoformes de NudCD1 observés en ARNm, l'expression protéique de celles-ci a été également mesurée chez les mêmes cellules par immunobuvardage (Figure 10). L'isoforme 1 est présente fortement chez les lignées cancéreuses et à des niveaux plus faibles chez les lignées non-cancéreuses. Pour leur part, NudCD1-2 et NudCD1-3 ne sont pas du tout exprimées au niveau protéique chez ces lignées cellulaires.

Ces expérimentations confirment les résultats antérieurs sur l'isoforme 1, démontrant un fort niveau d'expression de cette isoforme dans les lignées cancéreuses. Ces deux expériences ont toutefois permis d'observer un niveau d'expression plus faible, mais présent, chez les lignées cellulaires HIEC et CCD-18Co, respectivement des lignées épithéliales de l'intestin grêle et des fibroblastes du côlon, toutes deux non-tumorales. L'étude antérieure sur l'expression n'a montré qu'une expression testiculaire et cardiaque sur des tissus sains, et n'a pas montré d'expression chez les 26 autres types de tissus, incluant l'intestin grêle et le côlon (Yang *et al.*, 2001). Il est donc possible que NudCD1-1 soit exprimée dans d'autres tissus sains et que la protéine ait un rôle à jouer dans la croissance cellulaire normale. Il serait ainsi intéressant de mesurer son niveau d'expression dans d'autres types de tissus sains. Les données sur l'isoforme 2 viennent appuyer les observations précédentes sur son faible niveau d'expression dans certains cancers, toutefois en ne mesurant aucun niveau d'expression, ni même faible, chez les lignées cancéreuses étudiées. Cette isoforme n'a également présenté aucune expression chez les lignées saines. L'isoforme 3 est absente au niveau protéique dans ces mêmes lignées cancéreuses et chez les HIEC et CCD-18Co, mais est toutefois présente à des niveaux plus faibles en ARNm. Il serait ainsi intéressant de vérifier dans de nouvelles lignées cellulaires tumorales afin de confirmer leur présence ou absence lors de cancers. NudCD1-2 et NudCD1-3 n'ont été vérifiées que dans très peu de lignées cellulaires saines. Afin d'approfondir les connaissances sur l'importance de ces isoformes, il faudrait vérifier leur expression dans un plus grand nombre de tissus sains, comme lors de l'étude de Yang *et al.* en 2001 sur

NudCD1-1, car elles n'ont peut-être simplement pas de rôle à jouer dans l'intestin ou le côlon. Une étude de surexpression protéique de ces deux isoformes dans des cellules non-cancéreuses comme les NIH 3T3, analysant la prolifération, la résistance à l'apoptose, l'invasion et autres phénomènes tumorigéniques comme lors de l'étude de Rao *et al.* permettrait de déterminer si ces isoformes ont la capacité de causer le cancer.

#### **4.2 Les isoformes 2 et 3 sont dégradées rapidement par le protéasome par rapport à l'isoforme 1**

En raison de l'absence des différentes isoformes dans les lignées cellulaires utilisées, en plus de permettre les techniques d'immunoprécipitation utilisées ultérieurement, des lignées cellulaires exprimant des constructions stables des isoformes de NudCD1 avec une étiquette GFP sur les extrémités N et C-terminales dans des lignées cellulaires d'ostéosarcome U-2 OS ont été produites (Figure 11). On note que les expressions des isoformes 2 et 3 sont nettement plus faibles en comparaison à la première isoforme, malgré la concentration semblable de doxycycline utilisée, ainsi que le même vecteur d'expression, particulièrement dans le cas des constructions utilisant le vecteur pgLAP5.2 où la tétracycline présente dans le milieu de culture semble suffire pour induire la synthèse de l'isoforme 1 chez les cellules contrôles non-induites. On note également qu'il y a expression dans le cas de l'isoforme 1 dans le pgLAP5.2. Cette expression est due à la présence de tétracycline dans le FBS utilisé en culture cellulaire. La faible quantité de tétracycline est capable d'induire une synthèse de protéine recombinante et considérant la forte expression ou stabilité de cette isoforme, on note sa présence malgré l'absence de doxycycline pour l'induire. Il est possible que cette faible expression soit due à une régulation post-traductionnelle de ces deux isoformes sur leur région N-terminale spécifique. Les lignées U-2 OS pgLAP5.2 ont donc été utilisées pour vérifier la stabilité des isoformes *in vivo*. Des traitements sur ces lignées au MG132, un inhibiteur du protéasome, ont montré une augmentation de la concentration protéique des isoformes 2 et 3 par rapport aux cellules induites non traitées (Figure 13). Les cellules exprimant l'isoforme 1 n'ont pas montré de différences d'expression par rapport au contrôle. Ce résultat démontre que les isoformes 2 et 3 sont moins stables que NudCD1-1 et que leur faible expression par rapport à celle-ci pourrait être en partie attribuée à une dégradation rapide par le protéasome. Cette

dégradation pourrait également participer à l'explication de l'absence totale d'expression protéique de l'isoforme 3 chez les lignées cellulaires étudiées malgré l'expression d'ARNm observée en RT-PCR, impliquant qu'il est possible que cette isoforme soit exprimée, mais avec un temps de demi-vie très inférieur à NudCD1-1.

#### **4.3 Les 3 isoformes de NudCD1 sont localisées à différents endroits de la cellule.**

Les constructions cellulaires exprimant les différentes NudCD1 étiquetées GFP ont été utilisées pour déterminer si les différences structurales entre les isoformes pouvaient influencer leur localisation cellulaire. De plus, pour prévenir l'influence de l'ajout d'une étiquette GFP et ainsi vérifier si l'emplacement de l'étiquette pouvait interférer avec la destination de la protéine, les expériences d'immunofluorescence ont été faites en utilisant les constructions pgLAP1 et pgLAP5.2, exprimant l'étiquette GFP en N et en C-terminal respectivement. Les expériences avec l'étiquette en N-terminal montrent que la localisation cellulaire de l'isoforme 1 est majoritairement nucléaire et faiblement cytoplasmique (Figure 14). L'isoforme 2 est présente principalement au niveau du cytosol et faiblement au niveau nucléaire. Pour sa part, la troisième isoforme semble se situer autant au niveau du noyau que du cytosol. Les mêmes expériences ont été faites avec l'étiquette en C-terminal pour montrer encore une fois que l'isoforme 1 est principalement nucléaire, que l'isoforme 2 est très majoritairement située au cytosol alors que la troisième isoforme est pancellulaire (Figure 15). La présence de l'étiquette en N ou en C-terminal n'influence donc pas ou que très peu la localisation cellulaire des différentes isoformes, principalement au niveau de NudCD1-2 où l'étiquette en C-terminal a montré moins de localisation nucléaire que lorsqu'elle est placée en N-terminal. La troisième isoforme en particulier semble, en plus de sa localisation cytoplasmique et nucléaire, être présente dans un compartiment spécifique. La proximité au noyau porte à croire que celui-ci pourrait se retrouver impliqué dans le réticulum endoplasmique, mais une co-immunofluorescence avec des marqueurs des compartiments cellulaires n'a pas été faite, alors il est présentement impossible de confirmer sa localisation spécifique. Des expériences faites et rendu disponible sur le Human Protein Atlas démontrent également la présence de la protéine endogène au niveau du noyau et du cytosol (Berglung *et al.*, 2008). Des expériences de localisation par

fractionnements cellulaires ont également été effectuées, mais n'ont pas donné de résultats concluants. Il serait ainsi très intéressant de refaire des immunofluorescences pour NudCD1-3 en utilisant des marqueurs de différent compartiments pour déterminer précisément l'emplacement de cette isoforme, afin de donner de meilleures idées sur son rôle potentiel dans la cellule. Les isoformes de NudCD1 ont donc des localisations différentes dans la cellule et celle-ci pourrait influencer leurs partenaires d'interaction protéiques ainsi que leur rôle individuel.

#### **4.4 Partenaires d'interaction des différentes isoformes de NudCD1**

En utilisant les constructions cellulaires U-2 OS NudCD1-GFP pgLAP5.2 et d'autres constructions, les U-2 OS NudCD1-BirA pgLAP5.2, des expériences de spectrométrie de masse en tandem ont été effectuées pour identifier des partenaires d'interaction des 3 isoformes. Les constructions NudCD1-BirA ont été préalablement testées avec ou sans induction à la doxycycline et avec ajout de biotine pour permettre la biotinylation des protéines adjacentes aux isoformes de NudCD1 dans la cellule (Figure 13). Les deux méthodes, GFP et BirA sont complémentaires et permettent d'identifier autant des protéines ayant des liens directs lors de l'immunoprécipitation avec GFP, que des protéines ayant été à proximité des protéines étiquetées BirA. L'attrait d'utiliser la protéine de fusion BirA est qu'elle permet de biotinyler les lysines des protéines interagissant et à proximité de la protéine d'intérêt. BirA forme un gradient de biotinylation autour de la protéine, biotinylant plus fortement les protéines à proximité que les protéines plus loin selon un gradient de concentration (Roux *et al.*, 2012). De cette manière, les protéines interagissant directement avec la cible vont être très fortement biotinylées par rapport aux protéines avoisinantes, tout en permettant d'identifier ces dernières pour donner une bonne idée des complexes impliqués. La méthode de biotinylation possède également l'avantage de permettre des lavages plus stringents lors des précipitations à affinité et permettent donc d'enlever plus de protéines non-spécifiques lors de l'analyse par spectrométrie de masse.

Lors des expériences de spectrométrie effectuées avec le pgLAP5.2 GFP (Figure 17) l'isoforme 1 n'a présenté que 3 protéines autres que NudCD1 possédant un ratio

d'enrichissement supérieur à 2,0 par rapport au contrôle (Tableau 4). L'analyse sur l'isoforme 2 n'a quant à elle identifié que 4 protéines alors que la troisième isoforme s'est associée à 8 protéines (Tableau 4). Trois protéines ont interagi conjointement avec les isoformes 2 et 3, alors qu'aucune des protéines n'ayant interagi avec l'isoforme 1 n'a été retrouvée enrichie avec les isoformes 2 ou 3 (Tableau 4).. Il est étonnant de trouver une quantité si faible de protéines ayant interagi avec l'isoforme 1 de NudCD1 considérant sa forte expression dans le vecteur pgLAP5.2. Les mêmes expériences ont été faites avec les constructions pgLAP1, mais celles-ci n'ont jamais montré d'enrichissement supérieur à 2,0 pour la protéine NudCD1, et ce sur les 3 isoformes, rendant ces expériences invalides. Les expériences avec la protéine de fusion BirA ont pour leur part montré davantage de protéines associées aux différentes isoformes de NudCD1 (Figure 18). Les analyses de l'isoforme 1 ont montré une interaction avec 23 protéines (Tableau 5). Les isoformes 2 et 3 ont montré l'enrichissement de 18 et de 19 protéines, respectivement, outre NudCD1 (Tableau 5). De ces protéines, une seule est partagée uniquement par les deux premières isoformes alors qu'une seule est partagée uniquement par les isoformes 1 et 3. Les trois isoformes ont toutefois en commun 7 protéines. Les isoformes 2 et 3 partagent pour leur part 7 protéines n'interagissant pas avec NudCD1-1.

Les analyses ont montré que mis à part NudCD1, une seule protéine a interagi avec un ratio suffisant autant lors des expériences utilisant une immunoprécipitation GFP qu'une précipitation par affinité avec le système de biotine ligase. Cette protéine est DHX15. Il est toutefois intéressant de noter qu'une des 3 protéines ayant interagi lors de l'expérience NudCD1-1-GFP, COPB1, est un membre du complexe COPI. Un autre membre du complexe COPI, COPG2, a interagi avec NudCD1 lors des expériences utilisant BirA. Une étude a précédemment identifié des protéines de la famille des DEAD/DEAH box hélicases en plus de différents membres du complexe COPI (Taipale *et al.*, 2014). Les protéines DEAD/DEAH jouent d'importants rôles dans la biogénèse des ribosomes ainsi que dans l'épissage du pré-ARNm, via des changements structuraux du spliceosome grâce à leur activité hélicase ou en dérangeant des interactions protéines-ARN (Abdelhaleem *et al.*, 2003, Liang et Cheng, 2015). DHX15, un membre de la famille des DEAD/DEAH, est un orthologue de Prp43 chez la levure, qui est impliquée dans le désassemblage du spliceosome après les étapes de transestérifications lors de l'épissage du pré-ARNm (Chen

*et al.*, 2013). Le complexe COPI est mieux connu comme étant responsable du transport rétrograde des protéines membranaires entre le réticulum endoplasmique et l'appareil de golgi (Beck *et al.*, 2009). Ce complexe possède également une importance dans la tumorigénèse, alors que les membres de ce complexe sont surexprimés dans plusieurs lignées cellulaires cancéreuses et que leur déplétion nuit à l'autophagie de manière à diminuer la survie cellulaire (Claerhout *et al.*, 2012). La troisième protéine ayant interagi lors de l'expérience GFP est GRWD1, protéine peu caractérisée à ce jour qui pourrait faciliter la fixation des protéines MCM à la chromatine lors de la réplication de l'ADN (Sugimoto *et al.*, 2015). Les protéines ayant interagi avec la première isoforme lors de l'expérience avec BirA sont majoritairement impliquées dans différents endroits de la cellule et ayant différentes fonctions, principalement au niveau de l'organisation du cytosquelette, du cycle cellulaire, de la transcription et de l'apoptose (Figure 19). Les protéines étant les plus enrichies, outre DHX15 et COPG2, sont PCNP, CTTN et EFHD2 (Tableau 5). PCNP, la protéine ayant le plus grand enrichissement, une protéine également peu étudiée jusqu'à présent, est suggérée comme pouvant potentiellement avoir un impact au niveau de la régulation du cycle cellulaire ou de la stabilité génomique (Mori *et al.*, 2004). CTTN jouerait pour sa part un rôle dans la réorganisation du cytosquelette d'actine (von Holleben *et al.*, 2011). EFHD2 est une protéine se liant à l'actine permettant entre autres l'adhésion cellulaire et la signalisation par BCR (Park *et al.*, 2016). La grande variété de protéines identifiées dans l'expérience BirA, ainsi que la variété de fonctions associées à ces protéines rendent difficile à identifier un rôle plus probable pour NudCD1-1 par cette expérience uniquement. L'isoforme 1 pourrait donc avoir plusieurs rôles potentiels dans la cellule. Elle pourrait agir au niveau de la transcription de certains gènes ou agir de manière à réguler le cycle cellulaire ou même l'apoptose. Ces protéines n'ont toutefois pas été identifiées également lors de la spectrométrie de masse avec l'étiquette GFP. Certaines protéines ont interagi avec les trois isoformes lors de l'expérience BirA, toutefois, en raison des résultats communs lors des deux expériences en plus de l'étude de Taipale *et al.*, l'isoforme 1 de NudCD1 semble interagir principalement avec la protéine DHX15 et des membres du complexe COPI. L'étude précédente identifiait toutefois les DEAD/DEAH box hélicases DHX8, DHX16, DHX36 et DHX40 comme interagissant avec NudCD1-1, alors que les expériences de spectrométrie, autant avec l'étiquette GFP que BirA, ne démontrent

qu'une forte interaction avec DHX15 et aucune autre DEAH/DEAD box hélicase. Considérant la localisation majoritairement nucléaire et faiblement cytoplasmique, il est possible que NudCD1-1 puisse interagir au niveau du complexe COPI dans le cytoplasme, alors qu'il possède un autre rôle dans le noyau, impliquant la protéine DHX15. Il est également possible que l'isoforme 1 partage des fonctions avec les deux autres isoformes via leur forte homologie.

L'isoforme 2 en GFP a présenté 4 partenaires d'interaction : PSMD2, FECH, LOXL1 et SRCRB4D (Tableau 4). Aucune de ces protéines n'est retrouvée lors de l'expérience BirA. Trois de ces protéines se sont également retrouvées enrichies lors de l'immunoprécipitation de l'isoforme 3 : PSMD2, LOXL1 et SRCRB4D (Tableau 4). La protéine PSMD2, une sous-unité protéosomale non-catalytique, est liée à l'activité du protéasome et une étude montre que de l'ARN interférence sur cette protéine mène à une diminution protéosomale, diminue la croissance cellulaire et induit l'apoptose chez des lignées de cellules tumorales pulmonaires (Matsuyama *et al.*, 2011). La protéine LOXL1 est connue pour agir de manière à stabiliser la forme fibrillaire de FBLN5 dans l'élastogénèse (Liu *et al.*, 2004). SRCRB4D de son côté fait partie d'une superfamille de récepteurs scavengers et pourrait être impliquée dans des fonctions homéostatique de base comme la défense immunitaire innée (Padilla *et al.*, 2002). La protéine uniquement identifiée dans les analyses sur l'isoforme 2, FECH, est mitochondriale et impliquée dans la biosynthèse de l'hème (Yoon et Cowan, 2004). Cette dernière protéine semble moins probable en raison de sa localisation mitochondriale, mais comme la localisation déterminée n'est que cytoplasmique et qu'aucune vérification compartimentale n'a été faite, il est possible que NudCD1-2 soit impliqué dans ce phénomène dans le foie et la moëlle osseuse (Kolesnikov *et al.*, 2015). Les 18 protéines identifiées lors de l'expérience utilisant la biotine ligase fusionnée ont été séparées selon 3 principales catégories fonctionnelles : des protéines de réponse aux chocs thermiques, des protéines agissant dans l'organisation du cytosquelette et des protéines agissant dans la transcription (Figure 19). Parmi ces protéines, celle ayant le ratio d'enrichissement le plus élevé est DNAJA2, une protéine co-chaperone de HSP70, impliquée dans la réponse aux chocs thermiques (Rauch et Gestwicki, 2014). La majorité des autres protéines ont été également identifiées chez NudCD1-1 et NudCD1-3. La seule

exclusive ayant un rôle différent est EIF2S1, impliquée dans la traduction ribosomale (Zoll *et al.*, 2002).

L'isoforme 3 ne partage pas non plus de partenaires d'interaction identifiés par les deux méthodes. En plus des 3 protéines également présentes dans les immunoprécipitations GFP avec la seconde isoforme, les protéines MED4, STK32B, EPPK1, MFSD9 et PIK3C2B ont été trouvées (Tableau 4). MED4 est une protéine du complexe médiateur impliqué dans la transcription (Rachez *et al.*, 1999). STK32B, une kinase peu étudiée pouvant interagir avec HSP90 (Taipale *et al.*, 2012). EPPK1 est suggérée comme étant associée avec la motilité cellulaire et contribue à l'architecture tissulaire via un réarrangement des filaments intermédiaires (Shimada *et al.*, 2013). MFSD9 est un membre de la superfamille major facilitator qui sont des transporteurs membranaires de petits solutés solubles (Pao *et al.*, 1998). PIK3C2B est un membre de la famille des kinases PIK3, impliquées dans la cascade de signalisation des EGFR (Arcaro *et al.*, 2000). Malgré ces partenaires communs, les protéines identifiées pour ces deux isoformes sont toutes impliquées à des niveaux différents et il est ainsi difficile de spéculer sur l'implication de ces isoformes sans l'analyse BirA.

Comme pour l'isoforme 2, NudCD1-3 a également interagi lors de l'expérience BirA avec des protéines agissant principalement au niveau de la réponse aux chocs thermiques, de la réorganisation cytosquelettique ou mouvement nucléaire et de la transcription (Figure 19). La protéine ayant le plus fort ratio d'enrichissement, FKBP4, n'est pas partagée chez les autres isoformes, mais est également impliquée au niveau de la réponse aux chocs thermiques (Peattie, 1992). NudCD1-1 partage également des partenaires avec ses isoformes. Parmi les 7 protéines communes, 5 d'entre-elles, MAP4, FLNA, TAGLN2, CCT8 et CORO1B sont impliquées à divers endroits du transport cellulaire, de la motilité et impliquées au niveau du cytosquelette en général (Kitazawa *et al.*, 2000, Cai *et al.*, 2005, Seo *et al.*, 2010, Adams *et al.*, 2012, Na *et al.*, 2015). Ces partenaires communs pourraient tendre vers une explication que la partie commune entre les isoformes leur permettrait d'avoir un impact au niveau du cytosquelette, possiblement d'une manière commune à NudC, NudCL et NudCL2. Parmi les protéines ayant interagi avec les isoformes 2 et 3 lors de l'expérience BirA, MKL2, EEF1D, RUVBL1 et RUBVL2 sont les partenaires ayant un



rôle dans la transcription (Selvaraj et Prywes, 2003, Kaitsuka *et al.*, 2011, Niewiarowski *et al.*, 2010). La majorité des protéines associées à la réponse aux chocs thermiques sont communes aux deux isoformes : STIP1, HSP90AB1, DNAJA1 et EEF1D. Elles sont impliquées dans la réponse aux chocs d'HSP90, notamment HSP90AB1 qui est la forme constitutivement produite d'HSP90, et d'HSP70 (Woodford *et al.*, 2016, Retzlaff *et al.*, 2009, Stark *et al.*, 2014). EEF1D pour sa part permet la transcription de gènes impliqués dans ces types de réponses (Kaitsuka *et al.*, 2011). Notamment, NudC est l'un des partenaires commun aux deux isoformes. Cette interaction pourrait indiquer un rôle commun entre NudCD1-2 et 3 ainsi que NudC. NudC possède un rôle dans différents processus, comme la mitose, le mouvement cellulaire et nucléaire et la réponse inflammatoire (Zhang *et al.*, 2002). Ces isoformes interagissent avec des protéines impliquées avec le cytosquelette, suggérant que ces isoformes pourraient avoir un rôle complémentaire à NudC dans le mouvement nucléaire. Les partenaires d'interaction identifiés permettent ainsi de suggérer que ces isoformes seraient impliquées dans la transcription de gène, le mouvement nucléaire ou au niveau de la réponse aux chocs thermiques. Considérant les 492 acides aminés communs entre les deux isoformes, il est très probable que les deux isoformes possèdent des rôles et partenaires d'interaction similaires dans la cellule. L'immunofluorescence sur la 2e isoforme montre notamment sa localisation majoritairement cytoplasmique, où se trouve également la chaperone HSP90, et appuierait la supposition d'un rôle cytoplasmique lors de la réponse aux chocs thermiques via HSP90 et possiblement HSP70. La troisième isoforme, pour sa part, semble se situer dans plusieurs compartiments de la cellule, dont le noyau et ce qui pourrait être le réticulum endoplasmique. Elle pourrait ainsi avoir un rôle semblable à l'isoforme 2 au sein du cytoplasme, et avoir un rôle transcriptionnel dans le noyau. Advenant que la protéine se situe également au réticulum endoplasmique, elle pourrait agir de concert avec HSP90B1 pour supporter son activité chaperone dans ce compartiment cellulaire (Christianson *et al.*, 2008). La différence en N-terminal entre les deux protéines pourrait impliquer un signal de localisation cellulaire différent et ne pas avoir d'impact au niveau des interactions ou de la potentielle activité chaperone des isoformes 2 et 3. Des prédictions informatiques de NLS démontrent que l'isoforme 1 aurait un signal fort situé au domaine N-terminal, pouvant peut-être expliquer la plus forte présence au noyau (Lin et Hu, 2013). Des études futures

sur ces isoformes devraient confirmer l'interaction avec certains partenaires d'interaction par co-immunoprécipitation. Il serait également intéressant de vérifier par immunofluorescence, couplée à la microscopie confocale, la compartimentation spécifique de NudCD1-3 et possiblement confirmer sa présence au réticulum endoplasmique. Ces isoformes ne sont toutefois pas exprimées dans les lignées cellulaires cancéreuses et non-cancéreuses étudiées, suggérant qu'elles n'ont pas de fonction spécifique dans le cancer. En considérant l'expression jusqu'à maintenant exclusivement testiculaire des isoformes 2 et 3, il est possible que celles-ci aient des partenaires spécifiques dans ce tissu et qui ne se trouvent pas dans le modèle cellulaire étudié, qui sont des cellules d'ostéosarcome, empêchant donc de bien saisir le rôle de ces protéines (Yan *et al.*, 2004, Gerhard *et al.*, 2004). Somme toute, en se basant principalement sur les résultats offerts par la spectrométrie de masse sur les isoformes 2 et 3 avec l'étiquette BirA, ces protéines semblent être impliquées dans la réponse aux chocs thermiques comme chaperones ou co-chaperones, le mouvement cellulaire et la transcription de gènes. Davantage de recherches devront être faites sur ces isoformes au niveau testiculaire afin de mieux comprendre la raison de leur forte expression dans ce tissu, du moins pour NudCD1-2. NudCD1-1 de son côté, semble également pouvoir être impliquée au niveau du cytosquelette, potentiellement comme chaperone, mais sa localisation principalement nucléaire, en plus de la confirmation d'un fort interacteur identifié lors des deux méthodes, laisse croire que son rôle principal se situerait avec ce partenaire au niveau du noyau. La présence du domaine NudC, en plus des partenaires d'interaction impliqués dans diverses activités chaperones, rendrait intéressant l'analyse de l'activité chaperone des trois isoformes de NudCD1. Une méthode pour analyser l'activité chaperone de ces isoformes serait de vérifier l'activité enzymatique de protéines connues, comme la citrate synthase, d'inactiver à la chaleur ces protéines, avant d'ajouter ou de ne pas ajouter les différentes isoformes aux protéines inactives. L'activité enzymatique, dans cas-ci, serait alors mesurée par spectrophotométrie pour voir si l'enzyme retrouve sa fonction avec les différentes isoformes de NudCD1(Hristozova *et al.*, 2016).

#### 4.5 L'isoforme 1 de NudCD1 interagit avec la protéine DHX15

En raison des insuccès à trouver des partenaires d'interaction communs lors des deux expériences de spectrométrie de masse pour les isoformes 2 et 3, de la faible expression de celles-ci, en plus de l'absence de ces isoformes dans les lignées cellulaires cancéreuses étudiées, le reste de cette étude s'est concentré sur l'interaction de l'isoforme 1 de NudCD1 avec la protéine DHX15. Cette interaction entre ces deux protéines a d'abord été confirmée *in vivo* par transfection transitoire et co-immunoprécipitation dans des cellules HEK 293T (Figure 20). Cette expérience a également montré que l'interaction de DHX15 ne se fait qu'avec NudCD1-1, alors qu'aucune liaison entre DHX15 et NudCD1-2 ou 3 n'a été observée. D'autres expériences de spectrométrie de masse en tandem ont été faites, en utilisant encore une fois les protéines de fusion GFP et BirA associées avec DHX15 (Figure 21). Ces expériences ont permis d'identifier 16 protéines interagissant par la méthode GFP et 58 par la méthode BirA (Tableau 6). L'interaction NudCD1-1 et DHX15 a été encore une fois confirmée lors des deux expériences alors que NudCD1 a été parmi les plus enrichies lors des deux méthodes. Il y a 4 (GFP) et 8 (BirA) protéines qui sont impliquées au niveau de la régulation de la transcription, par exemple via la modification d'histones (Figure 22). Toutefois, la majorité de ces protéines, soit 10 dans l'expérience GFP et 33 dans l'expérience BirA, sont impliquées dans des mécanismes d'épissage de pré-ARNm (Figure 22). Notamment, 10 protéines enrichies ont été retrouvées à hauts niveaux dans les deux expériences, soit les protéines NudCD1, PAXBP1, TFIP11, SF3A1, SF3B2, CHERP, U2SURP, RBM17, SF3B1 et SF3B3. Les protéines SF3A1, SF3B2, SF3B1, SF3B3, ainsi que la protéine SF3A3, identifiée uniquement lors de la précipitation GFP, font partie des complexes SF3A et SF3B. Ces complexes sont connus pour agir de concert avec U2AF lors de la stabilisation du snRNP U2 au SB lors de la formation du complexe A (Will et Luhrmann, 2011). La protéine RBM17, en plus des protéines RBM5, RBM10, RBM 25, RBM 26 et RBM27 retrouvées dans l'expérience BirA-DHX15, sont des protéines de liaison à l'ARN. RBM17 est connue comme étant impliquée dans l'épissage alternatif, notamment du gène FAS, impliqué dans la régulation apoptotique (Corsini *et al.*, 2007). RBM10 serait probablement également impliquée dans l'épissage (Inoue *et al.*, 2008). Les protéines RBM26 et 27 ne sont que peu caractérisées et n'ont pas de rôle potentiel connu. TFIP11 est l'homologue humain de Ntr1 chez la levure, qui forme un

complexe avec Ntr2, qui à son tour interagit avec Brr2 pour mener au recrutement de Prp43 au spliceosome pour permettre le désassemblage du spliceosome (Tsai *et al.*, 2007). Cette interaction confirmerait l'impact de DHX15 dans le désassemblage du spliceosome, comme pour Prp43 et Ntr1. U2SURP est une protéine associée à la snRNP U2 alors que CHERP est une protéine impliquée dans l'homéostasie du calcium dans le réticulum endoplasmique (O'Rourke *et al.*, 2003). La protéine RBM5 est impliquée dans l'épissage alternatif de gènes reliés à la régulation de l'apoptose et est connue comme interagissant avec DHX15 de manière à stimuler son activité hélicase (Niu *et al.*, 2012). PAXBP1 pour sa part se lie aux facteurs de transcriptions Pax3 et Pax7 pour les faire interagir avec des complexes méthyltransférases et ainsi régulerait la prolifération de cellules musculaires progénitrices par épigénétisme chez la souris (Diao *et al.*, 2012). Une autre méthode pour confirmer l'interaction de l'isoforme 1 de NudCD1 et DHX15 a été de vérifier la co-localisation de ces protéines par immunofluorescence (Figure 23). Les deux protéines se sont retrouvées dans le noyau cellulaire et possèdent des points de co-localisation. Certains de ces points, sinon tous, pourraient toutefois être des nuclear speckles, soit des réservoirs à protéines majoritairement impliquées lors de l'épissage (Spector et Lamond, 2011). Pour vérifier cette possibilité, des expériences subséquentes d'immunofluorescence de co-localisation entre ces protéines, en plus d'utiliser un anticorps contre SC35, un marqueur de nuclear speckles seront nécessaires. Une immunofluorescence de DHX15 avec un partenaire d'interaction identifié lors de l'analyse en spectrométrie de masse, SF3A3, a également été effectuée pour montrer la co-localisation avec un membre de l'épissage (Figure 23). Les partenaires d'interaction communs aux deux expériences de spectrométrie, en plus de ceux trouvés seulement lors de l'analyse GFP et l'analyse BirA, démontrent que DHX15, comme Prp43 chez la levure, serait impliquée au niveau de l'épissage de pré-ARNm et pourrait potentiellement agir dans des phénomènes de régulation entre la transcription et l'épissage, en plus d'interagir avec l'isoforme 1 de NudCD1.

Ces résultats montrent que la carcinogénèse induite par l'expression de NudCD1 dans les lignées cellulaires pourrait être dépendante de modifications lors de l'épissage, ainsi que possiblement la transcription, via son interaction avec la protéine DHX15. Ces deux protéines n'ont toutefois pas de partenaires d'interaction communs, supposant que le rôle l'interaction entre NudCD1-1 et DHX15 ne mènerait pas à la formation d'un complexe et

que NudCD1-1 ne serait pas directement impliquée au niveau du spliceosome. L'étude présente ne permet toutefois pas de déterminer le résultat de l'interaction entre ces deux protéines. La présence de motifs  $\beta$ -propeller ne permet pas de déterminer le mode d'action d'une protéine. Ces motifs sont variés et retrouvés dans une pléthore de protéines. Les motifs à 7 lames sont particulièrement reconnus comme ayant le plus de variétés fonctionnelles (Chen *et al.*, 2011). Il est possible que celui présent sur NudCD1-1 lui permette non seulement de reconnaître ses partenaires, comme DHX15, mais aussi d'effectuer des modifications sur ses cibles via une activité enzymatique impliquant ou non l'ajout d'un ligand, ou même l'activité chaperone de son domaine NudC. Il serait intéressant de voir si NudCD1 stimule ou inhibe l'activité hélicase de DHX15 par essai de déroulement d'ARNdb avec et sans la protéine NudCD1 *in vitro*. Des études démontrent que DHX15 jouerait un rôle important lors de la réponse immunitaire contre les infections virales, en contribuant à l'activation des voies NF- $\kappa$ B, JNK et p38 de manière à augmenter la productions de cytokines et de promouvoir l'apoptose (Mosallanejad *et al.*, 2014, Kaltschmidt *et al.*, 2000). L'augmentation de l'expression de DHX15 a également montré des effets anti-migratoires et anti-prolifératifs dans des lignées cellulaires cancéreuses gastriques (Xiao *et al.*, 2016). Neuf gènes ont une diminution de leur expression chez des cellules cancéreuses HeLa lorsque NudCD1 est sous-exprimée : CTSL, MMP15, uPAR, VEGF, COX-2, S100A4, MUC1, MDM2 et RAC1 (Wang *et al.*, 2008). Aucun de ces produits de gènes n'a été enrichi lors de la spectrométrie de masse sur NudCD1. L'expression de NudCD1 semblerait même être corrélée avec l'expression de la protéine MDM2, une E3 ubiquitine ligase impliquée dans l'ubiquitination d'IGF1R et de p53 (Rao *et al.*, 2014, Girnita *et al.*, 2003). L'action de MDM2, impliquant l'ubiquitination d'IGF1R et de  $\beta$ -arrestin1, permettrait la régulation positive de la signalisation d'IGF1R (Girnita *et al.*, 2007). MDM2 est principalement connue comme inhibant p53, qui est normalement impliquée dans l'activation d'arrêt cellulaire et d'apoptose (Moll et Petrenko, 2003). Il est possible que le contrôle de l'expression de MDM2 par NudCD1-1 soit médié via l'interaction avec DHX15 en modifiant son activité avant l'épissage ou son influence sur la transcription de gènes, spécifiquement les 9 gènes mentionnés précédemment. En plus de l'impact de MDM2 sur la signalisation IGF1R, il est possible qu'il y ait un impact sur le niveau de p53 et d'autres cibles telles que pRB, HIPK2, Chk2 et HDAC (Riley et Lozano,

2012). L'activation d'autres voies médiées par MDM2 serait toutefois, au mieux, complémentaire à celle d'IGF1R, alors que l'inhibition pharmacologique d'IGF1R chez des cellules surexprimant NudCD1 diminue significativement la croissance (Rao *et al.*, 2014).

Considérant les interactions fortes de DHX15 avec des protéines comme SF3A et SF3B, il est possible que DHX15 agisse aussi sur des événements de formation des premiers complexes du spliceosome. Il serait intéressant d'étudier l'interactome de DHX15 dans un cas de déplétion de NudCD1, pour voir la modulation des interactions de DHX15 avec et sans présence de NudCD1. Vérifier l'impact d'une diminution protéique de DHX15 par ARN interférence, ou en trouvant un inhibiteur pharmacologique spécifique, sur la prolifération, migration, la résistance à l'apoptose et autres phénomènes associés à la tumorigénèse passant par le récepteur IGF1R, permettrait de déterminer si l'action oncogène de NudCD1 passe par cette hélicase à l'ARN. Ces études seraient toutefois difficiles à effectuer si DHX15 est absolument nécessaire à l'épissage de pré-ARNm, car il serait impossible d'associer les résultats potentiels à l'interaction NudCD1-DHX15, mais plutôt à l'impact de la diminution de DHX15 sur l'épissage. Les analyses sur NudCD1-1 ont également montré l'interaction avec des membres du complexe COPI, impliqués dans la survie cellulaire (Claerhout *et al.*, 2012). Bien qu'il soit probable que l'action de l'isoforme 1 passe par ce complexe pour moduler négativement l'autophagie, le faible nombre de partenaires du complexe qui ont été identifiés en plus de la localisation majoritairement nucléaire de cette isoforme rend moins plausible cette possibilité. L'interaction de NudCD1 avec le complexe COPI pourrait être investiguée davantage pour voir si NudCD1 peut possiblement perturber la formation du complexe et jouer un rôle au niveau du trafic cellulaire lors de l'autophagie.

Cette étude dans son ensemble a comporté plusieurs forces et limitations. La technique de RT-PCR a permis d'identifier la présence des différentes isoformes dans les lignées cellulaires, toutefois sans la quantifier, ce qu'une technique comme la qRT-PCR aurait pu faire. Les lignées cellulaires choisies auraient eu avantage à être plus diversifiées, la majorité d'entre elles provenaient du même type de tissus cancéreux alors qu'aucune lignée cellulaire cancéreuse provenant de poumons, de sein ou de prostate, qui sont parmi les cancers les plus mortels, n'a été utilisée (Canada, 2016). Malgré l'utilisation du vecteur

pgLAP5.2 possédant un promoteur d'expression fort, l'expression des isoformes 2 et 3 a été faible lors de la majorité des expériences, incluant lors de l'utilisation de la protéine de fusion BirA, qui permet la biotinylation des protéines adjacentes qui est stable malgré la possible dégradation de ces isoformes. L'utilisation de MG132 pour vérifier la dégradation protéique ne considère que le protéasome et non la possible dégradation lysosomale. Les immunofluorescences effectuées n'ont pas permis de vérifier si les protéines étaient localisées à des compartiments spéciaux, comme le réticulum endoplasmique ou la mitochondrie. La situation aurait pu être corrigée en utilisant des anticorps marqueurs de ces compartiments. Les deux techniques de spectrométrie de masse sont complémentaires pour trouver des interactions plus directes ainsi que pour trouver des interactions plus faibles ou étant plus loin dans un complexe et qui sont perdues lors des lavages. L'utilisation de spectrométrie de masse en tandem couplée à la méthode de quantification SILAC est une des forces de cette étude, car elle permet d'identifier précisément et de quantifier un très grand nombre de protéines malgré un nombre limité de peptides analysés. Les techniques de spectrométrie de masse ne permettent toutefois pas de déterminer la nature ni l'effet des interactions. Des expériences de double-hybride ou de co-immunoprécipitation *in vitro* seraient nécessaires pour déterminer si ces interactions sont directes ou indirectes. Pour compléter cette étude, il serait intéressant d'inspecter les isoformes de NudCD1 au niveau de leur domaines et différences. La génération de mutants de NudCD1, telle que l'utilisation d'un mutant ne comprenant que le domaine NudC, ainsi qu'un mutant clivant une partie de la région N-terminale de NudCD1, permettrait d'observer l'effet qui serait commun entre les isoformes, pour ensuite refaire des études de spectrométrie de masse et observer les partenaires afin de mieux comprendre l'effet de leurs domaines spécifiques. Il serait également intéressant de produire des mutants de DHX15 perturbant spécifiquement l'interaction entre celle-ci et NudCD1 pour analyser sa fonction sans l'impact de cette dernière. Des phosphorylations sur NudCD1 et ses isoformes sont possibles (Zhou *et al.*, 2013). Il est probable qu'il y ait d'autres modifications post-traductionnelles, telle que l'ubiquitination, qui influencent l'activité de ces protéines, leur localisation ou leur dégradation. Connaître le mécanisme de dégradation ou comprendre comment le favoriser serait une avancée dans la thérapie visant les patients atteints d'un cancer surexprimant la protéine NudCD1.

## 5. CONCLUSION

La présente étude avait comme objectif d'approfondir les connaissances sur les 3 différentes isoformes produites par le gène NudCD1 en identifiant tout d'abord quelles isoformes étaient présentes dans un contexte de cancer, de déterminer la localisation cellulaire de ces trois isoformes, ainsi que de déterminer leurs fonctions dans le cancer en identifiant leurs partenaires d'interaction spécifiques. Les expériences ont montré que seule la première isoforme est exprimée dans les lignées cellulaires cancéreuses alors des ARNm de l'isoforme 3 étaient présents, mais aucune évidence n'est présente pour sa forme protéique. Les isoformes 2 et 3 sont dégradées rapidement par le protéasome par rapport à NudCD1-1, pouvant expliquer l'absence de la 3e isoforme chez les lignées cellulaires étudiées. Une étude plus exhaustive devra par contre être faite afin de confirmer l'absence de ces isoformes dans le cancer, en plus de vérifier leur expression dans les tissus sains.

Les isoformes ont des localisations cellulaires différentes leur permettant des interactions avec différentes protéines. L'isoforme 1 est majoritairement nucléaire, l'isoforme 2 est majoritairement cytoplasmique, alors que NudCD1-3 a une localisation pancellulaire. Les analyses de spectrométrie de masse ont montré que l'isoforme 1 interagissait avec l'ARN hélicase DHX15 au niveau du noyau de la cellule ainsi qu'avec des membres du complexe de transport vésiculaire rétrograde de l'appareil de Golgi vers le réticulum endoplasmique COPI. Les isoformes 2 et 3 ont plusieurs partenaires d'interaction communs impliqués dans différentes fonctions cellulaires importantes. Les partenaires d'interaction identifiés suggèrent que NudCD1-2 pourrait participer à la réponse aux chocs thermiques médiée par la chaperone HSP90 ainsi que le mouvement nucléaire, alors que NudCD1-3 pourrait également être impliquée dans ceux-ci, en plus d'avoir un rôle dans la transcription nucléaire. Les interactions comprenant ces deux isoformes n'ont toutefois pas été confirmées et ces rôles demeurent, à ce stade, hypothétiques. NudCD1-3 semble également avoir une compartimentation spécifique au niveau du cytoplasme et pourrait posséder un rôle spécifique à cet endroit. Des nouvelles expériences de microscopie confocales avec des marqueurs de compartiments pourraient aider à distinguer la localisation de NudCD1-3. Les expériences de spectrométrie de NudCD1-2 et NudCD1-3 devraient également être



refaites, possiblement dans un autre modèle cellulaire, tel que des cellules testiculaires pour comprendre leur fonction dans un tissu spécialisé.

Parmi les interactions identifiées, celle entre NudCD1-1 et DHX15 a été validée en partie par co-immunoprécipitation. La spectrométrie de masse sur DHX15 a confirmé la présence de cette protéine dans le mécanisme de l'épissage de pré-ARNm ainsi qu'avec des partenaires orchestrant lors de la transcription. DHX15 est également connue pour avoir un impact au niveau des voies de signalisation de NF- $\kappa$ B, p38 et JNK, favorisant l'apoptose (Mosallanejad *et al.*, 2014, Kaltschmidt *et al.*, 2000). NudCD1 régule positivement l'expression de la protéine MDM2, celle-ci régulant entre autres p53 et IGF1R pour favoriser la survie et la prolifération cellulaire. Les résultats de la présente étude démontrent que NudCD1-1 pourrait réguler les voies d'IGF1R et la tumorigénèse par son interaction avec DHX15, qui influencerait l'épissage ou la transcription de manière à permettre la plus grande expression de MDM2.

La première isoforme démontre un rôle important dans le cancer, il serait ainsi avantageux d'étudier plus en détail cette protéine en particulier. La production de lignées cellulaires où le gène de NudCD1 est délété, produite par utilisation du système CRISPR/Cas9, permettrait d'observer si le gène est nécessaire pour des phénomènes comme la prolifération ou la croissance. Puisque sa principale partenaire d'interaction, DHX15, est impliquée au niveau du spliceosome, une étude de RNA-Seq par Illumina permettrait de voir les modifications au niveau du transcriptôme. Celui-ci pourrait être analysé selon des conditions cancéreuses normales où NudCD1-1 est surexprimé, ainsi qu'en condition de délétion ou de sous-expression de NudCD1-1. Les résultats de cette analyse permettrait de mieux comprendre comment NudCD1 influence la tumorigénèse. Davantage d'analyses par spectrométrie de masse permettraient de vérifier l'ajout de modifications post-traductionnelles sur NudCD1 et ainsi mieux comprendre ses mécanismes d'activation et de dégradation. Connaître les mécanismes menant à la dégradation et ainsi l'inhibition de NudCD1 pourrait ouvrir des options thérapeutiques anti-cancéreuses, via des molécules promouvant l'inhibition de NudCD1-1.

L'identification des partenaires de DHX15 a permis de déterminer que NudCD1-1 agit potentiellement en modifiant l'épissage ou la transcription. L'analyse de l'interactome de

DHX15 en conditions de délétion du gène de NudCD1, ou en conditions de sous-expression, permettrait de comparer avec les résultats obtenus et cibler la modulation d'interaction avec des partenaires spécifiques. Ces découvertes permettraient d'en savoir plus sur l'influence de NudCD1-1 sur DHX15. Une sous-expression par ARNsh de DHX15 dans des cellules de lignées tumorales où NudCD1-1 est surexprimée identifierait si la modulation des phénotypes tumoraux passe par DHX15, ou par d'autres partenaires potentiels comme le complexe COPI ou d'autres voies non-identifiées.

Cette étude n'a pas permis d'identifier clairement de rôles pour les isoformes 2 et 3 de NudCD1. Des analyses de surexpression de ces deux isoformes dans des lignées cellulaires non-tumorales n'exprimant pas ces isoformes, telles que les NIH 3T3 utilisées dans l'étude de Rao *et al.*, permettraient d'observer des effets cellulaires comparables à ceux liés à la surexpression de NudCD1-1. Il serait aussi intéressant de voir si les trois isoformes ont une activité chaperone fonctionnelle, supposée en raison de leur domaine NudC. Considérant l'absence ou faible expression de NudCD1 dans les différents tissus sains de l'organisme, il serait souhaitable d'analyser l'expression de ce gène, plus particulièrement de la première isoforme, dans une banque de tissus cancéreux humains. NudCD1-1 pourrait potentiellement permettre d'évaluer à quel niveau d'avancement la tumeur se situe et permettre une meilleure décision lors du choix de traitement.

Somme toute, les résultats de cette étude ont ainsi permis de mieux comprendre l'impact cellulaire des différentes isoformes de NudCD1. Ils ont permis de mieux caractériser les isoformes 2 et 3 en plus de démontrer que celles-ci ne sont probablement pas impliquées dans le cancer. Mieux comprendre les mécanismes de signalisations et protéines impliqués par l'expression de NudCD1-1 dans le cancer permettrait d'identifier potentiellement des cibles intéressantes pour la synthèse de nouvelles molécules impliquées dans les thérapies anti-cancéreuses et, par le fait même, donner plus d'options dans le combat contre le cancer.

## 6. LISTE DES RÉFÉRENCES

- Abdelhaleem M. Do human RNA helicases have a role in cancer?. *Biochim Biophys Acta*. 2004;1704(1):37-46.
- Adams M, Simms RJ, Abdelhamed Z, et al. A meckelin-filamin A interaction mediates ciliogenesis. *Hum Mol Genet*. 2012;21(6):1272-86.
- Adelman K, Lis JT. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet*. 2012;13(10):720-31.
- Arcaro A, Zvelebil MJ, Wallasch C, Ullrich A, Waterfield MD, Domin J. Class II phosphoinositide 3-kinases are downstream targets of activated polypeptide growth factor receptors. *Mol Cell Biol*. 2000;20(11):3817-30.
- Aumais JP, Williams SN, Luo W, et al. Role for NudC, a dynein-associated nuclear movement protein, in mitosis and cytokinesis. *J Cell Sci*. 2003;116(Pt 10):1991-2003.
- Barboric M, Lenasi T, Chen H, Johansen EB, Guo S, Peterlin BM. 7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development. *Proc Natl Acad Sci USA*. 2009;106(19):7798-803.
- Beck R, Rawet M, Ravet M, Wieland FT, Cassel D. The COPI system: molecular mechanisms and function. *FEBS Lett*. 2009;583(17):2701-9.
- Berglund L, Björling E, Oksvold P, et al. A gene-centric Human Protein Atlas for expression profiles based on antibodies. *Mol Cell Proteomics*. 2008;7(10):2019-27. Human Protein Atlas disponible sur : [www.proteinatlas.org](http://www.proteinatlas.org)
- Breathnach R, Chambon P. Organization and expression of eucaryotic split genes coding for proteins. *Annu Rev Biochem*. 1981;50:349-83.
- Black DL. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem*. 2003;72:291-336.
- Blanchette M, Chabot B. Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. *EMBO J*. 1999;18(7):1939-52.
- Bourgeois CF, Popielarz M, Hildwein G, Stevenin J. Identification of a bidirectional splicing enhancer: differential involvement of SR proteins in 5' or 3' splice site activation. *Mol Cell Biol*. 1999;19(11):7347-56.
- Brooks CL, Gu W. p53 ubiquitination: Mdm2 and beyond. *Mol Cell*. 2006;21(3):307-15.
- Bucceri A, Kapitza K, Thoma F. Rapid accessibility of nucleosomal DNA in yeast on a second time scale. *EMBO J*. 2006;25(13):3123-32.
- Burotto M, Chiou VL, Lee JM, Kohn EC. The MAPK pathway across different malignancies: a new perspective. *Cancer*. 2014;120(22):3446-56.

- Cai L, Holoweckyj N, Schaller MD, Bear JE. Phosphorylation of coronin 1B by protein kinase C regulates interaction with Arp2/3 and cell motility. *J Biol Chem*. 2005;280(36):31913-23.
- Chen CK, Chan NL, Wang AH. The many blades of the  $\beta$ -propeller proteins: conserved but versatile. *Trends Biochem Sci*. 2011;36(10):553-61.
- Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol*. 2009;10(11):741-54.
- Chodavarapu RK, Feng S, Bernatavichute YV, et al. Relationship between nucleosome positioning and DNA methylation. *Nature*. 2010;466(7304):388-92.
- Christianson JC, Shaler TA, Tyler RE, Kopito RR. OS-9 and GRP94 deliver mutant alpha1-antitrypsin to the Hrd1-SEL1L ubiquitin ligase complex for ERAD. *Nat Cell Biol*. 2008;10(3):272-82.
- Claerhout S, Dutta B, Bossuyt W, et al. Abortive autophagy induces endoplasmic reticulum stress and cell death in cancer cells. *PLoS ONE*. 2012;7(6):e39400.
- Canada, Comité directeur des statistiques sur le cancer de la Société canadienne du cancer : Statistiques canadiennes sur le cancer 2016. Toronto (Ontario) : Société canadienne du cancer, 2016. Accessible à [cancer.ca/Canadian-Cancer-Statistics-2016-FR.pdf](http://cancer.ca/Canadian-Cancer-Statistics-2016-FR.pdf) (consulté en octobre 2016).
- Copps KD, White MF. Regulation of insulin sensitivity by serine/threonine phosphorylation of insulin receptor substrate proteins IRS1 and IRS2. *Diabetologia*. 2012;55(10):2565-82.
- Corsini L, Bonnal S, Bonna S, et al. U2AF-homology motif interactions are required for alternative splicing regulation by SPF45. *Nat Struct Mol Biol*. 2007;14(7):620-9.
- Crick F. Central dogma of molecular biology. *Nature*. 1970;227(5258):561-3.
- Dai C, Gu W. p53 post-translational modification: deregulated in tumorigenesis. *Trends Mol Med*. 2010;16(11):528-36.
- Das R, Yu J, Zhang Z, et al. SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol Cell*. 2007;26(6):867-81.
- David CJ, Manley JL. The search for alternative splicing regulators: new approaches offer a path to a splicing code. *Genes Dev*. 2008;22(3):279-85.
- Darnell JE. Reflections on the history of pre-mRNA processing and highlights of current knowledge: a unified picture. *RNA*. 2013;19(4):443-60.
- Deng L, Chen J, Zhong XR, et al. Correlation between activation of PI3K/AKT/mTOR pathway and prognosis of breast cancer in Chinese women. *PLoS ONE*. 2015;10(3):e0120511.

- Dhanasekaran N, Premkumar reddy E. Signaling by dual specificity kinases. *Oncogene*. 1998;17(11 Reviews):1447-55.
- Diao Y, Guo X, Li Y, et al. Pax3/7BP is a Pax7- and Pax3-binding protein that regulates the proliferation of muscle precursor cells by an epigenetic mechanism. *Cell Stem Cell*. 2012;11(2):231-41.
- Dong Chen, Waters SB, Holt KH, Pessin JE. SOS phosphorylation and disassociation of the Grb2-SOS complex by the ERK and JNK signaling pathways. *J Biol Chem*. 1996;271(11):6328-32.
- Egloff S, Murphy S. Cracking the RNA polymerase II CTD code. *Trends Genet*. 2008;24(6):280-8.
- Faircloth LM, Churchill PF, Caldwell GA, Caldwell KA. The microtubule-associated protein, NUD-1, exhibits chaperone activity in vitro. *Cell Stress Chaperones*. 2009;14(1):95-103.
- Fong YW, Zhou Q. Stimulatory effect of splicing factors on transcriptional elongation. *Nature*. 2001;414(6866):929-33.
- Fourmann JB, Schmitzová J, Christian H, et al. Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes Dev*. 2013;27(4):413-28.
- Fredericks AM, Cygan KJ, Brown BA, Fairbrother WG. RNA-Binding Proteins: Splicing Factors and Disease. *Biomolecules*. 2015;5(2):893-909.
- Fu Q, Wang W, Zhou T, Yang Y. Emerging roles of NudC family: from molecular regulation to clinical implications. *Sci China Life Sci*. 2016;59(5):455-62.
- Fürstenberger G, Senn HJ. Insulin-like growth factors and cancer. *Lancet Oncol*. 2002;3(5):298-302.
- Fülöp V, Jones DT. Beta propellers: structural rigidity and functional diversity. *Curr Opin Struct Biol*. 1999;9(6):715-21.
- Georgescu MM. PTEN Tumor Suppressor Network in PI3K-Akt Pathway Control. *Genes Cancer*. 2010;1(12):1170-7.
- Gerhard DS, Wagner L, Feingold EA, et al. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Res*. 2004;14(10B):2121-7.
- Geuens T, Bouhy D, Timmerman V. The hnRNP family: insights into their role in health and disease. *Hum Genet*. 2016;135(8):851-67.
- Girnita L, Girnita A, Larsson O. Mdm2-dependent ubiquitination and degradation of the insulin-like growth factor 1 receptor. *Proc Natl Acad Sci USA*. 2003;100(14):8247-52.

- Girnita L, Shenoy SK, Sehat B, et al.  $\beta$ -Arrestin is crucial for ubiquitination and down-regulation of the insulin-like growth factor-1 receptor by acting as adaptor for the MDM2 E3 ligase. *J Biol Chem*. 2005;280(26):24412-9.
- Girnita L, Shenoy SK, Sehat B, et al. Beta-arrestin and Mdm2 mediate IGF1 receptor-stimulated ERK activation and cell cycle progression. *J Biol Chem*. 2007;282(15):11329-38.
- Goodrich JA, Tjian R. Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell*. 1994;77(1):145-56.
- Haisa M. The type 1 insulin-like growth factor receptor signalling system and targeted tyrosine kinase inhibition in cancer. *J Int Med Res*. 2013;41(2):253-64.
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-74.
- Hartmann TB, Mattern E, Wiedemann N, et al. Identification of selectively expressed genes and antigens in CTCL. *Exp Dermatol*. 2008;17(4):324-34.
- Hastings ML, Wilson CM, Munroe SH. A purine-rich intronic element enhances alternative splicing of thyroid hormone receptor mRNA. *RNA*. 2001;7(6):859-74.
- Hirose Y, Tacke R, Manley JL. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes Dev*. 1999;13(10):1234-9.
- Horowitz DS. The mechanism of the second step of pre-mRNA splicing. *Wiley Interdiscip Rev RNA*. 2012;3(3):331-50.
- Hristozova N, Tompa P, Kovacs D. A Novel Method for Assessing the Chaperone Activity of Proteins. *PLoS ONE*. 2016;11(8):e0161970.
- Huang X, Li X, Ma Q, et al. Chronic alcohol exposure exacerbates inflammation and triggers pancreatic acinar-to-ductal metaplasia through PI3K/Akt/IKK. *Int J Mol Med*. 2015;35(3):653-63.
- Huang Y, Li W, Yao X, et al. Mediator complex regulates alternative mRNA processing via the MED23 subunit. *Mol Cell*. 2012;45(4):459-69.
- Inoue A, Tsugawa K, Tokunaga K, et al. S1-1 nuclear domains: characterization and dynamics as a function of transcriptional activity. *Biol Cell*. 2008;100(9):523-35.
- Izaurrealde E, Lewis J, McGuigan C, Jankowska M, Darzynkiewicz E, Mattaj IW. A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell*. 1994;78(4):657-68.
- Jurado AR, Tan D, Jiao X, Kiledjian M, Tong L. Structure and function of pre-mRNA 5'-end capping quality control and 3'-end processing. *Biochemistry*. 2014;53(12):1882-98.
- Kaitsuka T, Tomizawa K, Matsushita M. Transformation of eEF1B $\delta$  into heat-shock response transcription factor by alternative splicing. *EMBO Rep*. 2011;12(7):673-81.

- Kaltschmidt B, Kaltschmidt C, Hofmann TG, Hehner SP, Dröge W, Schmitz ML. The pro- or anti-apoptotic function of NF-kappaB is determined by the nature of the apoptotic stimulus. *Eur J Biochem.* 2000;267(12):3828-35.
- Kent OA, Ritchie DB, Macmillan AM. Characterization of a U2AF-independent commitment complex (E') in the mammalian spliceosome assembly pathway. *Mol Cell Biol.* 2005;25(1):233-40.
- Keren H, Lev-Maor G, Ast G. Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet.* 2010;11(5):345-55.
- Kitazawa H, Iida J, Uchida A, et al. Ser787 in the proline-rich region of human MAP4 is a critical phosphorylation site that reduces its activity to promote tubulin polymerization. *Cell Struct Funct.* 2000;25(1):33-9.
- Kohtz JD, Jamison SF, Will CL, et al. Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature.* 1994;368(6467):119-24.
- Kolesnikov SI, Popova AS, Krupitskaya LI, Sinitskii AI, Kolesnikova LI. Activity of Heme Synthesis Enzymes in the Bone Marrow and Liver of August and Wistar Rats During the Neonatal Period and After Acute Postnatal Hypoxia. *Bull Exp Biol Med.* 2015;160(2):193-5.
- Kornblihtt AR, De La Mata M, Fededa JP, Munoz MJ, Nogues G. Multiple links between transcription and splicing. *RNA.* 2004;10(10):1489-98.
- Kuemmerle JF. IGF-I elicits growth of human intestinal smooth muscle cells by activation of PI3K, PDK-1, and p70S6 kinase. *Am J Physiol Gastrointest Liver Physiol.* 2003;284(3):G411-22.
- Kwek KY, Murphy S, Furger A, et al. U1 snRNA associates with TFIIF and regulates transcriptional initiation. *Nat Struct Biol.* 2002;9(11):800-5.
- Lai MC, Lin RI, Tarn WY. Transportin-SR2 mediates nuclear import of phosphorylated SR proteins. *Proc Natl Acad Sci USA.* 2001;98(18):10154-9.
- Lane D, Levine A. p53 Research: the past thirty years and the next thirty years. *Cold Spring Harb Perspect Biol.* 2010;2(12):a000893.
- Lardelli RM, Thompson JX, Yates JR, Stevens SW. Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. *RNA.* 2010;16(3):516-28.
- Lenasi T, Barboric M. Mutual relationships between transcription and pre-mRNA processing in the synthesis of mRNA. *Wiley Interdiscip Rev RNA.* 2013;4(2):139-54.
- Li B, Carey M, Workman JL. The role of chromatin during transcription. *Cell.* 2007;128(4):707-19.

- Li R, Pourpak A, Morris SW. Inhibition of the insulin-like growth factor-1 receptor (IGF1R) tyrosine kinase as a novel cancer therapy approach. *J Med Chem.* 2009;52(16):4981-5004.
- Liang WW, Cheng SC. A novel mechanism for Prp5 function in prespliceosome formation and proofreading the branch site sequence. *Genes Dev.* 2015;29(1):81-93.
- Lin JR, Hu J. SeqNLS: nuclear localization signal prediction based on frequent pattern mining and linear motif scoring. *PLoS ONE.* 2013;8(10):e76864.
- Liu X, Zhao Y, Gao J, et al. Elastic fiber homeostasis requires lysyl oxidase-like 1 protein. *Nat Genet.* 2004;36(2):178-82.
- Long JC, Cáceres JF. The SR protein family of splicing factors: master regulators of gene expression. *Biochem J.* 2009;417(1):15-27.
- Longman D, McGarvey T, McCracken S, Johnstone IL, Blencowe BJ, Cáceres JF. Multiple interactions between SRm160 and SR family proteins in enhancer-dependent splicing and development of *C. elegans*. *Curr Biol.* 2001;11(24):1923-33.
- Mandal SS, Chu C, Wada T, Handa H, Shatkin AJ, Reinberg D. Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. *Proc Natl Acad Sci USA.* 2004;101(20):7572-7.
- Martini M, De santis MC, Braccini L, Gulluni F, Hirsch E. PI3K/AKT signaling pathway and cancer: an updated review. *Ann Med.* 2014;46(6):372-83. Huang *et al.*, 2015-25573338
- Massagué J, Czech MP. The subunit structures of two distinct receptors for insulin-like growth factors I and II and their relationship to the insulin receptor. *J Biol Chem.* 1982;257(9):5038-45.
- Matera AG, Wang Z. A day in the life of the spliceosome. *Nat Rev Mol Cell Biol.* 2014;15(2):108-21.
- Meek DW. Regulation of the p53 response and its relationship to cancer. *Biochem J.* 2015;469(3):325-46.
- Moll UM, Petrenko O. The MDM2-p53 interaction. *Mol Cancer Res.* 2003;1(14):1001-8.
- Monami G, Emiliozzi V, Morrione A. Grb10/Nedd4-mediated multiubiquitination of the insulin-like growth factor receptor regulates receptor internalization. *J Cell Physiol.* 2008;216(2):426-37.
- Mori T, Li Y, Hata H, Kochi H. NIRF is a ubiquitin ligase that is capable of ubiquitinating PCNP, a PEST-containing nuclear protein. *FEBS Lett.* 2004;557(1-3):209-14.
- Morris NR. Nuclear migration. From fungi to the mammalian brain. *J Cell Biol.* 2000;148(6):1097-101.
- Morrison DK. MAP kinase pathways. *Cold Spring Harb Perspect Biol.* 2012;4(11)



- Mosallanejad K, Sekine Y, Ishikura-kinoshita S, et al. The DEAH-box RNA helicase DHX15 activates NF- $\kappa$ B and MAPK signaling downstream of MAVS during antiviral responses. *Sci Signal*. 2014;7(323):ra40.
- Muller PA, Vousden KH. Mutant p53 in cancer: new functions and therapeutic opportunities. *Cancer Cell*. 2014;25(3):304-17.
- Na BR, Kim HR, Piragyte I, et al. TAGLN2 regulates T cell activation by stabilizing the actin cytoskeleton at the immunological synapse. *J Cell Biol*. 2015;209(1):143-62.
- Nechaev S, Adelman K. Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation. *Biochim Biophys Acta*. 2011;1809(1):34-45.
- Niewiarowski A, Bradley AS, Gor J, McKay AR, Perkins SJ, Tsaneva IR. Oligomeric assembly and interactions within the human RuvB-like RuvBL1 and RuvBL2 complexes. *Biochem J*. 2010;429(1):113-25.
- Niu Z, Jin W, Zhang L, Li X. Tumor suppressor RBM5 directly interacts with the DExD/H-box protein DHX15 and stimulates its helicase activity. *FEBS Lett*. 2012;586(7):977-83.
- O'Rourke FA, Laplante JM, Feinstein MB. Antisense-mediated loss of calcium homeostasis endoplasmic reticulum protein (CHERP; ERROT213-21) impairs Ca<sup>2+</sup> mobilization, nuclear factor of activated T-cells (NFAT) activation and cell proliferation in Jurkat T-lymphocytes. *Biochem J*. 2003;373(Pt 1):133-43.
- Orphanides G, Lagrange T, Reinberg D. The general transcription factors of RNA polymerase II. *Genes Dev*. 1996;10(21):2657-83.
- Packham S, Warsito D, Lin Y, et al. Nuclear translocation of IGF-1R via p150(Glued) and an importin- $\beta$ /RanBP2-dependent pathway in cancer cells. *Oncogene*. 2015;34(17):2227-38.
- Padilla O, Pujana MA, López-de la iglesia A, et al. Cloning of S4D-SRCRB, a new soluble member of the group B scavenger receptor cysteine-rich family (SRCR-SF) mapping to human chromosome 7q11.23. *Immunogenetics*. 2002;54(9):621-34.
- Pao SS, Paulsen IT, Saier MH. Major facilitator superfamily. *Microbiol Mol Biol Rev*. 1998;62(1):1-34.
- Park HJ, Park KH, Shin KS, Lee MK. The roles of cyclic AMP-ERK-Bad signaling pathways on 6-hydroxydopamine-induced cell survival and death in PC12 cells. *Toxicol In Vitro*. 2013;27(8):2233-41.
- Park KR, An JY, Kang JY, et al. Structural mechanism underlying regulation of human EFhd2/Swiprosin-1 actin-bundling activity by Ser183 phosphorylation. *Biochem Biophys Res Commun*. 2016;

- Peattie DA, Harding MW, Fleming MA, et al. Expression and characterization of human FKBP52, an immunophilin that associates with the 90-kDa heat shock protein and is a component of steroid receptor complexes. *Proc Natl Acad Sci USA*. 1992;89(22):10974-8.
- Perales R, Bentley D. "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol Cell*. 2009;36(2):178-91.
- Perrini S, Laviola L, Carreira MC, Cignarelli A, Natalicchio A, Giorgino F. The GH/IGF1 axis and signaling pathways in the muscle and bone: mechanisms underlying age-related skeletal muscle wasting and osteoporosis. *J Endocrinol*. 2010;205(3):201-10.
- Pons T, Gómez R, Chinae G, Valencia A. Beta-propellers: associated functions and their role in human diseases. *Curr Med Chem*. 2003;10(6):505-24.
- Proudfoot NJ. Ending the message: poly(A) signals then and now. *Genes Dev*. 2011;25(17):1770-82.
- Pullikuth AK, Catling AD. Scaffold mediated regulation of MAPK signaling and cytoskeletal dynamics: a perspective. *Cell Signal*. 2007;19(8):1621-32.
- Rachez C, Lemon BD, Suldan Z, et al. Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex. *Nature*. 1999;398(6730):824-8.
- Rao W, Xie G, Zhang Y, et al. OVA66, a tumor associated protein, induces oncogenic transformation of NIH3T3 cells. *PLoS ONE*. 2014;9(3):e85705.
- Rao W, Li H, Song F, et al. OVA66 increases cell growth, invasion and survival via regulation of IGF1R-MAPK signaling in human cancer cells. *Carcinogenesis*. 2014;35(7):1573-81.
- Rauch JN, Gestwicki JE. Binding of human nucleotide exchange factors to heat shock protein 70 (Hsp70) generates functionally distinct complexes in vitro. *J Biol Chem*. 2014;289(3):1402-14.
- Retzlaff M, Stahl M, Eberl HC, et al. Hsp90 is regulated by a switch point in the C-terminal domain. *EMBO Rep*. 2009;10(10):1147-53.
- Riera J, Lazo PS. The mammalian NudC-like genes: a family with functions other than regulating nuclear distribution. *Cell Mol Life Sci*. 2009;66(14):2383-90.
- Riley T, Sontag E, Chen P, Levine A. Transcriptional control of human p53-regulated genes. *Nat Rev Mol Cell Biol*. 2008;9(5):402-12.
- Riley MF, Lozano G. The Many Faces of MDM2 Binding Partners. *Genes Cancer*. 2012;3(3-4):226-39.
- Roberts AJ, Kon T, Knight PJ, Sutoh K, Burgess SA. Functions and mechanics of dynein motor proteins. *Nat Rev Mol Cell Biol*. 2013;14(11):713-26.
- Robinson MJ, Cobb MH. Mitogen-activated protein kinase pathways. *Curr Opin Cell Biol*. 1997;9(2):180-6.

Roux KJ, Kim DI, Raida M, Burke B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol.* 2012;196(6):801-10.

Schroer TA. New insights into the interaction of cytoplasmic dynein with the actin-related protein, Arp1. *J Cell Biol.* 1994;127(1):1-4.

Schwer B. A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol Cell.* 2008;30(6):743-54.

Selvaraj A, Prywes R. Megakaryoblastic leukemia-1/2, a transcriptional co-activator of serum response factor, is required for skeletal myogenic differentiation. *J Biol Chem.* 2003;278(43):41977-87.

Seo S, Baye LM, Schulz NP, et al. BBS6, BBS10, and BBS12 form a complex with CCT/TRiC family chaperonins and mediate BBSome assembly. *Proc Natl Acad Sci USA.* 2010;107(4):1488-93.

Shimada H, Nambu-niibori A, Wilson-morifuji M, et al. Epiplakin modifies the motility of the HeLa cells and accumulates at the outer surfaces of 3-D cell clusters. *J Dermatol.* 2013;40(4):249-58.

Shukla S, Kavak E, Gregory M, et al. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature.* 2011;479(7371):74-9.

Silverman EJ, Maeda A, Wei J, Smith P, Beggs JD, Lin RJ. Interaction between a G-patch protein and a spliceosomal DEXD/H-box ATPase that is critical for splicing. *Mol Cell Biol.* 2004;24(23):10101-10.

Smith CW, Nadal-ginard B. Mutually exclusive splicing of alpha-tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. *Cell.* 1989;56(5):749-58.

Smith CW, Chu TT, Nadal-ginard B. Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns. *Mol Cell Biol.* 1993;13(8):4939-52.

Stark H, Lührmann R. Cryo-electron microscopy of spliceosomal components. *Annu Rev Biophys Biomol Struct.* 2006;35:435-57.

Stark JL, Mehla K, Chaika N, et al. Structure and function of human DnaJ homologue subfamily a member 1 (DNAJA1) and its relationship to pancreatic cancer. *Biochemistry.* 2014;53(8):1360-72.

Suemori K, Fujiwara H, Ochi T, et al. Identification of an epitope derived from CML66, a novel tumor-associated antigen expressed broadly in human leukemia, recognized by human leukocyte antigen-A\*2402-restricted cytotoxic T lymphocytes. *Cancer Sci.* 2008;99(7):1414-9.

Suemori K, Fujiwara H, Ochi T, et al. Identification of a novel epitope derived from CML66 that is recognized by anti-leukaemia cytotoxic T lymphocytes. *Br J Haematol.* 2009;146(1):115-8.

- Sugimoto N, Maehara K, Yoshida K, et al. Cdt1-binding protein GRWD1 is a novel histone-binding protein that facilitates MCM loading through its influence on chromatin architecture. *Nucleic Acids Res.* 2015;43(12):5898-911.
- Sun Y, Liu WZ, Liu T, Feng X, Yang N, Zhou HF. Signaling pathway of MAPK/ERK in cell proliferation, differentiation, migration, senescence and apoptosis. *J Recept Signal Transduct Res.* 2015;35(6):600-4.
- Sundaram MV. Canonical RTK-Ras-ERK signaling and related alternative pathways. *WormBook.* 2013;:1-38.
- Taipale M, Krykbaeva I, Koeva M, et al. Quantitative analysis of HSP90-client interactions reveals principles of substrate recognition. *Cell.* 2012;150(5):987-1001.
- Taipale M, Tucker G, Peng J, et al. A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell.* 2014;158(2):434-48.
- Tange TO, Damgaard CK, Guth S, Valcárcel J, Kjems J. The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. *EMBO J.* 2001;20(20):5748-58.
- Teodoro JG, Evans SK, Green MR. Inhibition of tumor angiogenesis by p53: a new role for the guardian of the genome. *J Mol Med.* 2007;85(11):1175-86.
- Tsai RT, Tseng CK, Lee PJ, et al. Dynamic interactions of Ntr1-Ntr2 with Prp43 and with U5 govern the recruitment of Prp43 to mediate spliceosome disassembly. *Mol Cell Biol.* 2007;27(23):8027-37.
- Ullrich A, Gray A, Tam AW, et al. Insulin-like growth factor I receptor primary structure: comparison with insulin receptor suggests structural determinants that define functional specificity. *EMBO J.* 1986;5(10):2503-12.
- Vallee RB, Mckenney RJ, Ori-mckenney KM. Multiple modes of cytoplasmic dynein regulation. *Nat Cell Biol.* 2012;14(3):224-30.
- van der Feltz C, Anthony K, Brilot A, Pomeranz krummel DA. Architecture of the spliceosome. *Biochemistry.* 2012;51(16):3321-33.
- Von Holleben M, Gohla A, Janssen KP, Iritani BM, Beer-hammer S. Immunoinhibitory adapter protein Src homology domain 3 lymphocyte protein 2 (SLy2) regulates actin dynamics and B cell spreading. *J Biol Chem.* 2011;286(15):13489-501.
- Vousden KH, Lane DP. p53 in health and disease. *Nat Rev Mol Cell Biol.* 2007;8(4):275-83.
- Wada T, Penninger JM. Mitogen-activated protein kinases in apoptosis regulation. *Oncogene.* 2004;23(16):2838-49.
- Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell.* 2009;136(4):701-18.

Walker RA, Sheetz MP. Cytoplasmic microtubule-associated motors. *Annu Rev Biochem.* 1993;62:429-51.

Wang Q, Li M, Wang Y, et al. RNA interference targeting CML66, a novel tumor antigen, inhibits proliferation, invasion and metastasis of HeLa cells. *Cancer Lett.* 2008;269(1):127-38.

Warsito D, Sjöström S, Andersson S, Larsson O, Sehat B. Nuclear IGF1R is a transcriptional co-activator of LEF1/TCF. *EMBO Rep.* 2012;13(3):244-50.

Wick MJ, Dong LQ, Riojas RA, Ramos FJ, Liu F. Mechanism of phosphorylation of protein kinase B/Akt by a constitutively active 3-phosphoinositide-dependent protein kinase-1. *J Biol Chem.* 2000;275(51):40400-6.

Will CL, Lührmann R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol.* 2011;3(7)

Woodford MR, Dunn DM, Blanden AR, et al. The FNIP co-chaperones decelerate the Hsp90 chaperone cycle and enhance drug binding. *Nat Commun.* 2016;7:12037.

Wu CJ, Yang XF, McLaughlin S, et al. Detection of a potent humoral response associated with immune-induced remission of chronic myelogenous leukemia. *J Clin Invest.* 2000;106(5):705-14.

Wybenga-Groot LE, McGlade CJ. RTK SLAP down: the emerging role of Src-like adaptor protein as a key player in receptor tyrosine kinase signaling. *Cell Signal.* 2015;27(2):267-74.

Xiao YF, Li JM, Wang SM, et al. Cerium oxide nanoparticles inhibit the migration and proliferation of gastric cancer by increasing DHX15 expression. *Int J Nanomedicine.* 2016;11:3023-34.

Yan Y, Phan L, Yang F, et al. A novel mechanism of alternative promoter and splicing regulates the epitope generation of tumor antigen CML66-L. *J Immunol.* 2004;172(1):651-60.

Yang XF, Wu CJ, McLaughlin S, et al. CML66, a broadly immunogenic tumor antigen, elicits a humoral immune response associated with remission of chronic myelogenous leukemia. *Proc Natl Acad Sci USA.* 2001;98(13):7492-7. Wang *et al.*, 2008- 18534745

Yang Y, Yan X, Cai Y, Lu Y, Si J, Zhou T. NudC-like protein 2 regulates the LIS1/dynein pathway by stabilizing LIS1 with Hsp90. *Proc Natl Acad Sci USA.* 2010;107(8):3499-504.

Yoon T, Cowan JA. Frataxin-mediated iron delivery to ferrochelatase in the final step of heme biosynthesis. *J Biol Chem.* 2004;279(25):25943-6.

Yudkovsky N, Ranish JA, Hahn S. A transcription reinitiation intermediate that is stabilized by activator. *Nature.* 2000;408(6809):225-9.

Zeng C, Berget SM. Participation of the C-terminal domain of RNA polymerase II in exon definition during pre-mRNA splicing. *Mol Cell Biol*. 2000;20(21):8290-301.

Zhang C, Zhang W, Lu Y, et al. NudC regulates actin dynamics and ciliogenesis by stabilizing cofilin 1. *Cell Res*. 2016;26(2):239-53. Hartmann *et al.*, 2008- 17979976

Zhang MY, Huang NN, Clawson GA, et al. Involvement of the fungal nuclear migration gene nudC human homolog in cell proliferation and mitotic spindle formation. *Exp Cell Res*. 2002;273(1):73-84.

Zhang Y, Smith AD, Renfrow MB, Schneider DA. The RNA polymerase-associated factor 1 complex (Paf1C) directly increases the elongation rate of RNA polymerase I and is required for efficient regulation of rRNA synthesis. *J Biol Chem*. 2010;285(19):14152-9.

Zhang Y, Madl T, Bagdiul I, et al. Structure, phosphorylation and U2AF65 binding of the N-terminal domain of splicing factor 1 during 3'-splice site recognition. *Nucleic Acids Res*. 2013;41(2):1343-54.

Zhang Y, Goodfellow R, Li Y, et al. NEDD4 ubiquitin ligase is a putative oncogene in endometrial cancer that activates IGF-1R/PI3K/Akt signaling. *Gynecol Oncol*. 2015;139(1):127-33.

Zheng M, Cierpicki T, Burdette AJ, et al. Structural features and chaperone activity of the NudC protein family. *J Mol Biol*. 2011;409(5):722-41.

Zheng WH, Kar S, Quirion R. Insulin-like growth factor-1-induced phosphorylation of transcription factor FKHRL1 is mediated by phosphatidylinositol 3-kinase/Akt kinase and role of this pathway in insulin-like growth factor-1-induced survival of cultured hippocampal neurons. *Mol Pharmacol*. 2002;62(2):225-33.

Zhou H, Di palma S, Preisinger C, et al. Toward a comprehensive characterization of a human cancer cell phosphoproteome. *J Proteome Res*. 2013;12(1):260-71.

Zhou T, Zimmerman W, Liu X, Erikson RL. A mammalian NudC-like protein essential for dynein stability and cell viability. *Proc Natl Acad Sci USA*. 2006;103(24):9039-44.

Zhou Z, Licklider LJ, Gygi SP, Reed R. Comprehensive proteomic analysis of the human spliceosome. *Nature*. 2002;419(6903):182-5.

Zhu J, Mayeda A, Krainer AR. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol Cell*. 2001;8(6):1351-61.

Zhu XJ, Liu X, Jin Q, Cai Y, Yang Y, Zhou T. The L279P mutation of nuclear distribution gene C (NudC) influences its chaperone activity and lissencephaly protein 1 (LIS1) stability. *J Biol Chem*. 2010;285(39):29903-10.

Zoll WL, Horton LE, Komar AA, Hensold JO, Merrick WC. Characterization of mammalian eIF2A and identification of the yeast homolog. *J Biol Chem*. 2002;277(40):37079-87.

## 7. ANNEXE

### 7.1 Protein interaction network of alternatively spliced NudCD1 isoforms

# Protein interaction network of alternatively spliced NudCD1 isoforms.

*Patrick Asselin-Mullen, Anaïs Chauvin, Marie-Line Dubois, Romain Drissi, Dominique Lévesque and François-Michel Boisvert<sup>1</sup>*

<sup>1</sup> Corresponding author: François-Michel Boisvert

Department of Anatomy and Cell Biology, Université de Sherbrooke, 3201 Jean-Mignault, Sherbrooke, Québec, J1E 4K8, Canada

Phone: 819-821-8000 Ext.75430

Fax: 819-820-6831

Email: fm.boisvert@usherbrooke.ca

KEYWORDS: NudCD1, mRNA splicing, Mass Spectrometry, SILAC, Subcellular proteomics

## ABSTRACT

NudCD1, also known as CML66 or OVA66, is a protein initially identified as overexpressed in patients with chronic myelogenous leukemia. The mRNA of NudCD1 is expressed in heart and testis of normal tissues, and is overexpressed in several cancers. Previous studies have shown that the expression level of the protein correlates with tumoral phenotype, possibly interacting upstream of the Insulin Growth Factor - 1 Receptor (IGF-1R). The gene encoding the NudCD1 protein consists of 12 exons that can be alternative spliced, leading to the expression of three different isoforms. These isoforms possess a common region of 492 amino acids in their C-terminus region and have an isoform specific N-terminus. To determine the distinct function of each isoforms, we have localised the isoforms within the cells using immunofluorescence microscopy and used a quantitative proteomics approach (SILAC) to identify specific protein interaction partners for each isoforms. Localization studies showed a different subcellular distribution for the different isoforms, with the first isoform being nuclear, while the other two isoforms have distinct cytoplasmic and nuclear location. We found that the different NudCD1 isoforms have unique interacting partners, with the first isoform binding to a putative RNA helicase named DHX15 involved in mRNA splicing.



## INTRODUCTION

The nuclear distribution gene C (NudC) protein family is composed of four conserved proteins: NudC, NudC-like (NudCL), NudC-like 2 (NudCL2) and NudC domain containing 1 (NudCD1)<sup>1</sup>, the later also called chronic myelogenous leukaemia 66 (CML66)<sup>2</sup> or Ovarian cancer-associated antigen 66 (OVA66)<sup>3</sup>. These proteins share a conserved p23 domain conferring them a chaperone activity for binding to p23 and/or Heat shock protein 90 (Hsp90) client proteins<sup>4</sup>. It has been showed that NudC proteins play multiple roles in cell cycle progression, neuronal migration, inflammatory response, platelet production, carcinogenesis<sup>5-8</sup> and their expression is generally higher in proliferating cells<sup>9</sup>. Among this family, NudCD1/CML66 is the more distant family member and has the least characterized mechanism of action. NudCD1 is a tumour associated antigen highly expressed in human leukaemia, some solid tumours and tumour cell lines<sup>2,10</sup>. Alternative splicing (Figure 1A) of the mRNA results in three different isoforms sharing a common C terminus (66 kDa isoform 1 [583aa], 64 kDa isoform 2 [554 aa]<sup>2</sup>, and 61 kDa isoform. While these proteins are often expressed in different cancer cells and tumors, their expression in normal tissues is restricted to testis<sup>10,11</sup>. It also has been demonstrated that NudCD1 was broadly immunogenic, notably following the discovery of specific antibody in 18 to 38% of sera from patients with lung, melanoma and prostate cancers<sup>12,13</sup>.

Knock-down of NudCD1 results in an inhibition of cell proliferation, migration and invasion through regulation of the IGF-1R-MAPK pathway<sup>10,12</sup>, underlining the potential as a target for immunotherapeutic approaches in a variety of solid tumours. Using a high throughput assay to characterize the chaperone-cochaperone interaction network in human cells, Taipale et al. established that proteins from the NudC family specifically associate with structurally related  $\beta$ -propeller folds which include proteins such as Coat Protein I (COPI) complex proteins, DEAD/DEAH box helicases and proteins with other functions<sup>14</sup>. COPI complex proteins are involved in cellular trafficking, and modulate autophagy and cell death in cancer cells<sup>15</sup> and DEAD/DEAH box helicases are involved in almost all aspects of nucleic acid transactions<sup>16</sup>. Furthermore, NudCD1 was found associated with proteins involved in the dissociation of U snRNPs from the lariat-intro complex in experiments identifying protein interactions with the human Pre-mRNA processing 43 (hPrp43) and TFIP11, the human ortholog of yeast Prp43<sup>17</sup>, a post-splicing factor<sup>18</sup>.

Alternative splicing of pre-mRNAs is very common in humans and throughout most eukaryotes, occurring in up to 60% of all genes<sup>19</sup>. Indeed, as many as 95% of human genes containing several exons are undergoing alternative splicing<sup>20</sup>, and this level of tissue and functional diversity in higher organisms is thought to be due in part to alternative splicing<sup>21</sup>. Although alternative splicing is generally thought to be responsible for the diversity of gene products expressed from the genome, the complexity of alternative splicing at the proteome level remains to be characterised<sup>22</sup>. Alternative splicing is well documented at the transcript level, but large-scale proteomics experiments have mostly used a one gene, one protein approach to simplify the number of proteins to be analyzed. Most large scale experiments have focused on using antibodies recognizing a region common to the different isoforms, or have chosen the most characterized protein isoform to include in their study of identifying protein interaction or expression. In this article, we characterized the

different NudCD1 isoforms, identifying different roles and subcellular distribution for each of the three isoforms.

## EXPERIMENTAL PROCEDURES

**Cell and SILAC culture.** U2OS and U2OSFlp-In T-Rex (U2OS FT) cells were grown as adherent cells in Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal bovine serum, 100 U/ml penicillin/streptomycin and 2 mM GlutaMax. Additionally, U2OS Flp-In cells are maintained with 100 µg/ml Zeocin and 15 µg/ml Blasticidine-HCl. For culture in SILAC media, DMEM without arginine and lysine (Life Technologies A14431-01) was supplemented with 10% dialyzed fetal bovine serum (Invitrogen, 26400-044), 100 U/ml Penicillin/streptomycin and 2 mM GlutaMax. Isotopic arginine and lysine were added in either light (Arg0, Sigma, A5006; Lys0, Sigma, L5501), medium (Arg6, Cambridge Isotope Lab (CIL), CNM-2265; Lys4, CIL, DLM-2640), or heavy (Arg10, CIL, CNLM-539; Lys8, CIL, CNLM-291) to a 28 µg/ml and 49 µg/ml final concentration of arginine and lysine, respectively. To avoid proline to arginine conversion, an excess concentration of 10 µg/ml of L-proline was added to each medium.

**Gateway Technology and Stable Cell Lines Production.** NudCD1 isoforms 1, 2 and 3, as well as DHX15, were amplified by PCR using oligonucleotides with the BP recombination sites *attB* from a cDNA library generated by RT-PCR using an oligo-dT from mRNA isolated from U2OS cells by Trizol (Invitrogen). The BP recombination reaction using the BP Clonase™ (Life Technologies) was realized between the *attB*-containing PCR products and the *attP*-containing donor vector pDONR 221 (Life Technologies) to generate an entry clone. Then, LR recombination reaction was allowed by LR Clonase™ (Life Technologies) between *attL*-containing pDONR 221-NudCD1 isoforms 1, 2 or 3 and *attR*-containing destination vector pGLAP5.2-GFP. This destination vector is a modified version of pGLAP1<sup>23</sup> to generate an expression clone with GFP at the C-terminal. DHX15 was cloned into a modified pGLAP1 vector with a Myc tag at the N-terminus. U2OS FT stable cell lines expressing inducible GFP-tagged proteins were generated by transfecting pGLAP5.2-GFP plasmids containing the cDNA of interest along with pOG44 (the plasmid expressing the Flp-recombinase) into U2OS FT cells to allow integration at a specific genomic location<sup>24</sup>. U2OS cells were then selected with the addition of 100 µg/ml Hygromycin B and 10 µg/ml Blasticidine-HCl, and protein induction was achieved by adding 2 µg/ml Doxycyclin for 48 hours.

**Antibodies.** The following primary antibodies were used for immunoblotting and immunofluorescence experiments: anti-GFP (mouse monoclonal, Roche #11814460001), anti-GAPDH (rabbit monoclonal, Cell Signalling #14C10), anti-Myc (rabbit, Cell Signalling #71D10) and anti-SF3A3 (mouse monoclonal, Abcam #56823). The following secondary antibodies were used: anti-mouse IgG-HRP (goat polyclonal, Santa Cruz #sc-2005), anti-rabbit IgG-HRP (goat polyclonal, Santa Cruz #sc-2004), Alexa Fluor® 546 goat anti-mouse IgG (Invitrogen #1010044), Alexa Fluor® 546 goat anti-rabbit IgG (Invitrogen #949214), Alexa Fluor® 488 goat anti-mouse IgG (Invitrogen #1008801), Alexa Fluor® 488 goat anti-rabbit IgG (Invitrogen #1024116) and goat polyclonal anti-chicken IgY DyLight® 650 (Abcam #96954).

**Immunofluorescence Microscopy.** U2OS FT pgLAP5.2 NudCD1-GFP isoform1, 2 or 3 were cultured on glass coverslips in six-well plates. 50% confluent cells were then washed twice with PBS and fixed with 1% paraformaldehyde in PBS for 10 min at room temperature. Fixed cells were washed with PBS and permeabilized using 0.5% Triton X-100 in PBS for 10 min. After another wash in PBS, coverslips were incubated with primary antibodies diluted in PBS (rabbit or mouse antibody depending on the target protein) for 1 h at room temperature. Coverslips were washed once in 0.1% Triton X-100 in PBS and twice in PBS. Primary antibodies were detected with Alexa Fluor 488 goat anti-rabbit and/or Alexa Fluor 546 goat anti-mouse diluted in PBS for 1 h at room temperature. Coverslips were washed in 0.1% Triton X-100 in PBS containing 1  $\mu$ g/ml 4,6-diamidino-2-phenylindole (DAPI) allowing DNA counterstain for 5 min. After wash with PBS, coverslips were mounted on glass slides and dried overnight at 4°C in the darkness.

**Immunoprecipitation from SILAC Labeled Cells.** U2OS FT cells expressing isoform 1, 2 or 3 of NudCD1 were harvested separately by scraping in PBS and pellets were lysed in immunoprecipitation (IP) buffer (1% Triton X-100, 10 mM Tris pH 7.4, 150 mM NaCl) for 10 min on ice. Insoluble material was pelleted by centrifugation for 10 min at 13,000 g at 4°C and supernatants from the three SILAC conditions were combined. Equal amount of proteins were incubated with GFP-trap agarose beads (ChromaTek) for 2 h at 4°C. Beads were washed with IP buffer then with PBS. Finally they were resuspended in Laemmli sample buffer prior to SDS-PAGE.

**Gel electrophoresis and in-gel digestion.** Proteins were reduced in 10 mM DTT, alkylated in 50 mM iodoacetamide and incubated at 95°C for 5 min in 1X Laemmli buffer. They were then separated by one-dimensional SDS-PAGE (4–12% Bis-Tris Novex mini-gel, Life Technologies) and visualized by Coomassie staining (Simply Blue Safe Stain, Life Technologies). Following extensive washes in water, the gel was cut into slices and subjected to in-gel digestion with 12.5 ng/ml trypsin (Promega) in 20 mM  $\text{NH}_4\text{HCO}_3$ . Tryptic peptides were extracted by 1% formic acid, then 100% acetonitrile. Solvent was removed by lyophilization in a speed vacuum centrifuge, and the tryptic peptides were resuspended in 1% formic acid.

**LC-MS/MS.** Trypsin digested peptides were loaded and separated onto a Dionex Ultimate 3000 nanoHPLC system. 10  $\mu$ l of the sample (2  $\mu$ g) resuspended in 1% (v/v) formic acid was loaded with a constant flow of 4  $\mu$ l/min onto a trap column (Acclaim PepMap100 C18 column, 0.3 mm id x 5 mm, Dionex Corporation, Sunnyvale, CA). Peptides were then eluted off and loaded onto a PepMap C18 nano column (75  $\mu$ m x 50 cm, Dionex Corporation) with a linear gradient of 5–35% solvent B (90% acetonitrile with 0.1% formic acid) over a 4 h gradient with a constant flow of 200 nl/min. Peptides were then electrosprayed into an Orbitrap QExactive mass spectrometer (ThermoFischer Scientific Inc.) by an EasySpray source. The spray voltage was 2.0 kV and the temperature of the analytical column was 40°C. Acquisition of the full scan MS survey spectra (m/z 350–1600) in profile mode was performed in the Orbitrap at a resolution of 70,000 using 1,000,000 ions. Peptides selected for fragmentation by collision-induced dissociation were based on the ten highest intensities for the peptide ions from the preview scan. Normalized collision energy was set at 35% and resolution was set at 17,500 for 50,000 ions. Filling times was set to a maximum of 250 ms for the full scans and 60 ms for the MS/MS scans.

All unassigned charge states as well as singly, 7 and 8 charged species for the precursor ions were rejected. Additionally, a dynamic exclusion list was set to retain up to 500 entries with a maximum retention time of 40 seconds using a 10 ppm mass window. To improve the mass accuracy of survey scans, the lock mass option was enabled. Data acquisition was done using Xcalibur version 2.2 SP1.48.

***Quantification and Bioinformatics Analysis.*** Protein identification and quantification were performed using the MaxQuant software package version 1.5.2.8 as described previously<sup>25</sup> with the protein database from UniProtKB (Homo sapiens, 16/07/2013, 88,354 entries). For protein identification, carbamidomethylation on cysteine was used as a fixed modification and methionine oxidation and protein N-terminal acetylation were used as variable modifications. The enzyme was set to trypsin, with no cleavages on lysine or arginine before a proline, and up to two miscleavages were allowed. The mass tolerance was set at a maximum of 7 ppm for the precursor ions and 20 ppm for the fragment ions. Re-quantification of selected isotopic patterns was allowed to obtain ratios of all SILAC pairs<sup>26</sup>. We set a threshold of 5% for the false discovery rate (FDR) based on the criteria that the number of forward hits identified from the database was at least 20-fold higher than the number proteins identified in a database containing reverse protein sequences. For protein quantification, we set a minimum of 2 peptides identified for each protein. The mass spectrometry data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository with the dataset identifier PXD005493.

## RESULTS

### ***NudCD1 isoforms cloning, expression and localization.***

There are three NudCD1 isoforms that are generated by alternative splicing of exons 1, 2 and 4 (Figure 1A). Isoform 2 is the only one to include exon 2 resulting in a different N-terminal, whereas isoform 3 includes exon, similarly to isoform 1, but also includes exon 4 which has an additional translation start site. This results in a smaller isoform with a shorter and different N-terminal (Figure 1A). Isoforms 1 and 2 do not include exon 4. Isoform 1 is specifically overexpressed in leukemia and a variety of solid tumor cell lines and is normally detected in testis and heart tissues<sup>2</sup>, isoform 2 is predominantly expressed in testis and weakly expressed in tumor cells<sup>27</sup> and isoform 3 was identified in a full-length mammalian cDNA project<sup>28</sup>. To determine the functions of each isoforms, we designed fusion proteins composed of NudCD1 (isoform 1, 2 or 3) with GFP at the C-terminus of the protein, since the differences between the isoforms are found at the N-terminus of NudCD1. Three tet-inducible stable cell lines were obtained: U2OS FT pgLAP5.2 NudCD1-GFP isoform 1, 2 and 3. We verified the three NudCD1-GFP isoforms expression in U2OS FT by inducing expression by the addition of 2 µg/ml of doxycycline for 24h (Figure 1B). Isoform 1 was strongly expressed compared with isoforms 2 and 3 (Figure 1B). To determine whether some of the proteins were unstable and rapidly degraded, we induced expression of each isoforms by the addition of doxycycline for 24 hours, and incubated for 8 h with 10 µM MG132, a proteasome inhibitor (Figure 1C). The expression of all isoforms was strongly increased following treatment with MG132. The subcellular localization of each isoform was determined by immunofluorescence microscopy experiments (Figure 2). Expression of each isoforms was induced in U2OS FT pgLAP5.2 NudCD1-GFP cells by the addition of doxycycline for 24 hours, and the cells

were fixed prior to immunofluorescence labelling using a GFP antibody, and the nucleus was visualized using DAPI to allow DNA coloration. Imaging of the different isoforms showed a different cellular distribution for each isoform with a predominantly nuclear accumulation of NudCD1 isoform 1 (Figure 2A-C), cytoplasmic and nuclear accumulation of NudCD1 isoform 2 (Figure 2D-F) and a mostly cytoplasmic localization of NudCD1 isoform 3 (Figure 2G-I). These results demonstrate that each isoform of NudCD1 have distinct subcellular distribution.

### ***NudCD1 isoform-specific interacting partners***

We investigated NudCD1 interacting partners using a combination of immunoprecipitation followed by quantitative mass spectrometry analysis using triple-labeling SILAC<sup>29</sup> to measure differences between the interactions for each isoform, as well as contaminant proteins. Cells not expressing NudCD1 were grown in “light” (L) media and served as a control for nonspecific binding. To identify proteins specifically interacting with each isoforms of NudCD1, cells expressing NudCD1 isoform 1 were grown in “medium” (M) media and cells expressing either NudCD1 isoform 2 or isoform 3 were grown in cells grown in “heavy” (H) media. Cells were cultured in SILAC media for at least 6 passages to ensure complete incorporation of the isotopic amino acids, and each experiment were repeated twice. Following lysis, total cell lysates are mixed and subjected to immunoprecipitation using the recombinant GFP-Trap\_A reagent<sup>30</sup> by combining a control, isoform 1 and isoform 2 (Figure 3A), or a control, isoform 1 and isoform 3 (Figure 3B). We identified 162 interacting partners for isoform 1, 53 for isoform 2 and 73 for isoform 3 (ratio IP over contaminant above 1.5, Figure 3A-C and supplementary table 1). By comparing the different isoforms, we found 10 interacting partners specific to isoforms 1 and 2 (Figure 3A, C), 27 specific to isoforms 1 and 3 (Figure 3B, C) and 6 specific to isoforms 2 and 3. Four interacting partners were found in common between the three isoforms (Figure 3C and supplementary Table 1): FBLN1, FBXO21, LDHB and SRCRB4D.

For each isoform, we classified their partners depending on their cellular compartment localisation (Fig. 4A-C) and biological processes in which they are implicated (Fig. 4D-F), using the gene ontology annotations enrichment tool DAVID<sup>31,32</sup>. A large proportion of NudCD1 isoform 1 interacting partners were nuclear (Figure 4A), consistent with its subcellular distribution (Figure 2). Isoform 1 interacting proteins were strongly implicated in biological processes related to mRNA, either RNA polymerase II transcription or mRNA processing (Figure 4D). A larger proportion of proteins interacting with isoform 2 were cytoplasmic, membrane bound or annotated for cell projection (Figure 4B). No biological processes appeared to be predominant for isoform 2 with only low scoring functions such as intracellular processes, mRNA processing and apoptosis). Isoform 3 interacted with nuclear proteins, but also included several mitochondrial proteins consistent with a more granular cytoplasmic localization (Figure 2 and 4C).

### ***NudCD1-DHX15 interaction***

Of all the proteins identified, DHX15 was found interacting with NudCD1 with the highest enrichment ratio over contaminants, intensity and number of peptides, indicating a strong, near stoichiometric interaction with NudCD1. Interestingly, this interaction was only with the isoform 1, and not isoforms 2 or 3. DHX15 (DEAH-Box Helicase 15) is a

putative ATP-dependent RNA helicase implicated in pre-mRNA splicing. To validate the interaction identified by mass spectrometry, co-immunoprecipitation (IP) experiments were realized in U2OS cell lines expressing NudCD1-GFP isoforms 1, 2 or 3 and a Myc-tagged DHX15 (Figure 5). We have first proceeded with a GFP immunoprecipitation (Figure 5A) and then the reciprocal immunoprecipitation with a Myc antibody (Figure 5B) and revealed the NudCD1-DHX15 co-IP with anti-Myc and anti-GFP antibodies. In both co-immunoprecipitations, we could only detect a co-IP between DHX15 and the isoform 1 of NudCD1. No interaction between NudCD1 isoform 2 or 3 and DHX15 was detected. This confirmed the identification of DHX15 as a NudCD1 isoform 1 specific interactor.

To determine whether NudCD1 and DHX15 are part of the same protein complex, DHX15 protein interactions were identified by immunoprecipitation experiments followed by mass spectrometry analysis (Figure 6A and supplementary Table 2). Control cells were grown in “light” (L) media and cells expressing GFP-DHX15 were grown in “medium” (M) media. Following lysis, total cell lysates are mixed and subjected to immunoprecipitation using the recombinant GFP-Trap\_A reagent as described earlier, by combining the control with the GFP-DHX15 cell lysates. Many interacting partners of DHX15 were identified, including several splicing factors such as RBM10, RBM17, SF3A1 and A3, SF3B1 to B5, GCFC2, TFIP11, SNRPA1 and B2 (Figure 6A, B). Of these proteins, most of them belonged to the splicing factor 3b protein complex (Figure 6B), confirming a role in mRNA splicing for DHX15 and suggesting that it may regulate the U2 snRNP binding to the branchpoint sequence in pre-mRNA<sup>33</sup>. While NudCD1 was identified as a DHX15 interacting partners, there are no other overlap between the proteins interacting with NudCD1 isoform 1 and DHX15 interacting proteins, suggesting an exclusive interaction between DHX15 and NudCD1.

Several splicing factors including SF3 proteins localize to nuclear speckles or interchromatin granule clusters<sup>34</sup>. To determine whether NudCD1 and DHX15 interaction was mutually exclusive to the interaction of DHX15 with splicing factors, we investigated whether NudCD1 isoform 1 and DHX15 colocalized in nuclear speckles within the nucleus by immunofluorescence microscopy (Figure 7). U2OS FT cells transfected with Myc-DHX15 and NudCD1-GFP constructions were fixed and revealed with GFP and Myc antibodies (Figure 7A-D) allowing the detection of NudCD1 and DHX15, respectively. Both proteins displayed a nuclear localization (Figure 7B&C). However, while DHX15 localized to structure typical of speckle domains (Figure 7C), NudCD1 was found mostly diffuse through the nucleus, and excluded from the nucleoli (Figure 7B). To confirm that the localization pattern of DHX15 was consistent with splicing speckles, we also performed immunofluorescence experiments on DHX15 and one of the identified splicing factor SF3A3 (Figure 7E-H). The Splicing factor 3A subunit 3 is essential for the formation of the mature 17S U2 snRNP and the pre-spliceosome assembly<sup>35</sup>. U2OS cells transfected with Myc-DHX15 was also labelled using a SF3A3 antibody (Figure 7F&G). Cells expressing DHX15 were found to have a similar localization in nuclear speckles as shown by the co-localization with SF3A3 (Figure 7H). This indicates that NudCD1 does not co-localize with DHX15 in nuclear speckles, consistent with the interaction with NudCD1 in the absence of splicing factors.

## DISCUSSION

NudCD1 was initially cloned from a chronic myelogenous leukaemia cDNA expression library and was found highly expressed in leukaemia and solid tumor cell lines. However, its expression in normal tissues is restricted mainly to testis and heart <sup>2</sup>. It was later found that the NudCD1 gene encodes for three different isoforms that share a common C-terminus (Figure 1A), and that isoforms 1 and 2 had differential expression resulting from transcriptional regulation: isoform 1 was predominantly expressed in testis and overexpressed in tumor cell lines while isoform 2 was weakly expressed in tumor cell lines <sup>27</sup>. In cells expressing inducible NudCD1 isoforms, we found that the isoform 1 was much more stable compared to isoforms 2 and 3, and that the protein levels increased after inhibition of the proteasome with MG132 (Figure 1), but not completely. Because all stable cell lines used a unique integration sites, the differences in expression could not be explained by differences at the transcriptional level. This demonstrates that other factors influence the expression and/or the stability of each isoform in U2OS, as well as in other cell lines such as HCT116 and 293 (data not shown).

Interestingly, each of the three isoforms has distinct cellular localization (Figure 2), which could explain the large difference in protein stability observed. Isoform 1 is mainly nuclear while isoforms 2 and 3 are found in the cytoplasm, but with a different distribution. Isoform 2 displays a more diffuse cytoplasmic distribution while isoform 3 is found around the nuclear membrane in a punctate pattern, reminiscent of endoplasmic reticulum or the Golgi apparatus. It is thus not surprising that each isoform was found to have few proteins in common in the immunoprecipitation experiments and mass spectrometry analysis (Figure 3C), indicating that each isoform is associated with different protein complexes. For isoform 1, the proteins identified were consistent with its nuclear localization, as most of the proteins identified were involved in RNA polymerase II transcription or mRNA processing (Figure 4D). For isoform 2, almost a half of the interacting partners identified are proteins found in the cytoplasm, as well as other cellular subcompartments (Fig. 4B). Isoform 3 was found interacting with nuclear proteins, but also with proteins from specific cellular subcompartments such as the mitochondria or focal adhesion points (Fig. 4C). Of all the proteins identified, four were common to the three isoforms (Table 1), consistent with the lack of overlap in cellular localization. Using mass spectrometry and high-throughput LUMIER assays to characterize the chaperone interaction network, Taipale *et al.* also showed that NudCD1 interacts with proteins from the COP complex involved in cellular trafficking, as well as several DEAD/DEAH box helicases<sup>14</sup>, but did not differentiate between the different NudCD1 isoforms. COP complex proteins are involved in cellular trafficking, and appear to modulate autophagy and cell death in cancer cells<sup>15</sup> and DEAD/DEAH box helicases are involved in almost all the nucleic acid transactions<sup>16</sup>.

In previous studies, it was demonstrated that NudCD1 could play a role in tumorigenesis by increasing cellular proliferation, invasion and metastasis. Downregulation of NudCD1 in tumor cells that showed a high expression of the protein resulted in a downregulation of some oncogenic genes such as CTSL, MMP15, uPAR, VEGF, COX-2, S100A4, MUC1, MDM2 and RAC1<sup>10</sup> and an increase of the resistance to 5-fluorouracil-induced apoptosis<sup>11</sup>. Secondly, NudCD1 could promote oncogenesis by enhancing activation of the IGF-1R-ERK1/2 signalling pathway, notably implicated in cell

proliferation and differentiation, via its regulation of MDM2 (ubiquitin ligase for IGF-1R)<sup>12</sup>. None of these proteins were found as interacting partners of NudCD1, suggesting that the regulation of the level of these proteins is not through interaction with NudCD1, but could be a consequence of the expression of NudCD1 on other cellular processes that would be involved in regulating the expression of these proteins, such as transcription or splicing.

Of all the proteins identified, DHX15 was found consistently with the highest number of peptides and intensity, suggesting a near stoichiometric association with NudCD1 (Table 1). Moreover, this interaction was only observed with the isoform 1 of NudCD1, indicating that the interaction is either occurring only in the nucleus, or that the small difference at the N-terminus of NudCD1 is the region involved in the interaction. DHX15 has been proposed to be involved in mRNA splicing<sup>36</sup>, which was confirmed in our mass spectrometry analysis that identified several splicing factors interacting with DHX15, mostly from the SF3 complex (Figure 6). DHX15 was identified as the human homolog of the yeast protein Prp43 based on sequence similarities, and could replace the function of Prp43 in a yeast-based splicing assay<sup>37</sup>. Prp43 is involved in spliceosome disassembly via its interaction with Ntr1 and Ntr2 in yeast, through interaction with the N-terminal G-patch domain of Ntr1, resulting in the stimulation of the helicase activity of Prp43.<sup>38-41</sup> Ntr1 is also called Tufelin Interacting Protein 11 (TFIP11) and was identified as an interacting partner of DHX15 in our experiments (Figure 6). Our experiments have showed co-localization between DHX15 and proteins from the SF3 splicing factor complex that are essential for the formation of the mature 17S U2 snRNP and pre-spliceosome assembly<sup>35</sup>, further confirming a role for DHX15 in splicing. What is interesting here is that we did not observe any co-localization between NudCD1 and DHX15, and we did not identify any common interacting partners except for the fact that these two proteins were found as the highest interacting proteins with each other. This suggests that the interaction is mutually exclusive with the function of DHX15 in mRNA splicing, suggesting a potential role for NudCD1 in regulating or sequestering DHX15 away from splicing factors.

Thus, our study showed a different subcellular distribution for the different isoforms of NudCD1, with the first isoform being nuclear, while the other two isoforms have distinct cytoplasmic and nuclear location. We found that the different NudCD1 isoforms have unique interacting partners, with the first isoform binding to a putative RNA helicase named DHX15 involved in mRNA splicing. These data provide a molecular basis to explain the different contradictory roles observed for NudCD1 in cell proliferation and tumorigenesis upon up or downregulation.

## **ACKNOWLEDGEMENTS**

This work was funded to F.M.B. by Natural Sciences and Engineering Research Council of Canada (grant number 418404). F.M.B. is also a FRQS Junior II scholar.

## **ADDITIONAL INFORMATION**

### **Competing financial interests**

The authors declare no competing financial interests.

### **Author contribution statement**



P.A.M. performed all the experiments presented in figures 1,3,4,5 and 6. M.L.D. and R.D. performed the microscopy images in figures 2 and 7. A.C. and F.M.B. wrote the main manuscript text and D.L. provided technical assistance. All authors reviewed the manuscript.

## FIGURE LEGENDS

### **Figure 1 –Alternative splicing results in three different NudCD1 isoforms.**

**A)** Schematic representation of the first 4 exons of NudCD1 and the resulting isoforms that differs in their N-terminus, while all isoforms include exons 5 to 12. Isoform 1 consists of exons 1 and 3, isoform 2 consists of exons 2 and 3, while the isoform 3 includes the exons 1, 3 and 4, but uses an initiation codon in the fourth exon, resulting in a smaller protein. **B)** Total protein lysates from U2OS FT cells with doxycycline-induced GFP-tagged NudCD1 isoforms were analyzed by Western blotting using a GFP antibody to confirm expression of the different isoforms following induction using doxycycline (lane 2, 4, 6). **C)** Total protein lysates from U2OS FT cells with doxycycline-induced GFP-tagged NudCD1 isoform treated or not with MG132 were analyzed by Western blotting using a GFP antibody.

### **Figure 2 –The different NudCD1 protein isoforms have distinct subcellular distribution.**

Fluorescence microscopy analysis of the U2OSFT cells expressing the respective GFP-fusion proteins was performed using an antibody to GFP (panels A, D and G), and the nuclei were stained with DAPI (panels B, E and H). NudCD1 Isoform 1 shows a strictly nuclear accumulation (A), NudCD1 isoform 2 shows both a cytoplasmic and nuclear (D) and a predominantly cytoplasmic accumulation is shown for NudCD1 isoform 3 (G).

### **Figure 3 –Identification of NudCD1 isoform-specific interacting partners.**

NudCD1 interacting partners were identified using a combination of immunoprecipitation followed by quantitative mass spectrometry analysis using triple-labeling SILAC. Cells not expressing NudCD1 were grown in “light” (L) media, cells expressing NudCD1 isoform 1 were grown in “medium” (M) media and cells expressing either NudCD1 isoform 2 or isoform 3 were grown in cells grown in “heavy” (H) media. Immunoprecipitation were performed using the recombinant GFP-Trap\_A reagent by combining a control, isoform 1 and isoform 2 (**3A**), or a control, isoform 1 and isoform 3 (**3B**). The overlap between the proteins identified as interactors for each isoform is displayed as a Venn diagram (**C**).

### **Figure 4 – Gene ontology annotations enrichment for each NudCD1 isoforms.**

The proteins interacting with each of the NudCD1 isoforms were analysed for gene ontology annotations enrichment for cellular compartments (**A-C**) and biological processes (**D-F**) using DAVID.

### **Figure 5 – DHX15 interacts with NudCD1 isoform 1 only.**

To validate the interaction identified by mass spectrometry, co-immunoprecipitation experiments were realized in U2OS FT cell lines expressing NudCD1-GFP isoforms 1, 2 or 3 and co-transfected with a Myc-tagged DHX15. A GFP immunoprecipitation (**A**) and then

the reciprocal immunoprecipitation with a Myc antibody (**B**) were performed and revealed with Myc and GFP antibodies.

**Figure 6 –DHX15 interacts with splicing factors from the SF3 complex.**

(A) Control cells were grown in “light” (L) media and cells expressing GFP-DHX15 were grown in “medium” (M) SILAC media. Following lysis, total cell lysates are mixed and subjected to immunoprecipitation using the recombinant GFP-Trap\_A reagent. The graph represents the proteins enriched in the immunoprecipitation with a  $\log_2$  of the enrichment ratio in the x-axis, and a  $\log_{10}$  of the intensity in the y-axis.(B) A functional protein association networks using STRING shows several known interactions between members of the splicing factor SF3 complex.

**Figure 7 – DHX15 colocalizes with SF3 proteins in nuclear speckles, but not NudCD1.**

U2OS cells transfected with Myc-DHX15 and NudCD1-GFP constructions were fixed and revealed with GFP and Myc antibodies (Figure 7A-D). To confirm that the localization pattern of DHX15 was consistent with splicing speckles, we also performed immunofluorescence experiments on DHX15 and one of the identified splicing factor, SF3A3 (Figure 7E-H). The nucleus were stained with DAPI.

**References**

- 1 Fu, Q., Wang, W., Zhou, T. & Yang, Y. Emerging roles of NudC family: from molecular regulation to clinical implications. *Sci China Life Sci***59**, 455-462, doi:10.1007/s11427-016-5029-2 (2016).
- 2 Yang, X. F. *et al.* CML66, a broadly immunogenic tumor antigen, elicits a humoral immune response associated with remission of chronic myelogenous leukemia. *Proc Natl Acad Sci U S A***98**, 7492-7497, doi:10.1073/pnas.131590998 (2001).
- 3 Jin, S. *et al.* Humoral immune responses against tumor-associated antigen OVA66 originally defined by serological analysis of recombinant cDNA expression libraries and its potentiality in cellular immunity. *Cancer Sci***99**, 1670-1678, doi:10.1111/j.1349-7006.2008.00860.x (2008).
- 4 Garcia-Ranea, J. A., Mirey, G., Camonis, J. & Valencia, A. p23 and HSP20/alpha-crystallin proteins define a conserved sequence domain present in other eukaryotic protein families. *FEBS Lett***529**, 162-167 (2002).
- 5 Hatakeyama, H. *et al.* Protein clusters associated with carcinogenesis, histological differentiation and nodal metastasis in esophageal cancer. *Proteomics***6**, 6300-6316, doi:10.1002/pmic.200600488 (2006).
- 6 Lin, S. H. *et al.* Inhibition of prostate tumor growth by overexpression of NudC, a microtubule motor-associated protein. *Oncogene***23**, 2499-2506, doi:10.1038/sj.onc.1207343 (2004).
- 7 Miller, B. A. *et al.* A homolog of the fungal nuclear migration gene nudC is involved in normal and malignant human hematopoiesis. *Exp Hematol***27**, 742-750 (1999).
- 8 Suzuki, S. O. *et al.* Expression patterns of LIS1, dynein and their interaction partners dynactin, NudE, NudEL and NudC in human gliomas suggest roles in invasion and proliferation. *Acta Neuropathol***113**, 591-599, doi:10.1007/s00401-006-0180-7 (2007).

- 9 Gocke, C. D., Osmani, S. A. & Miller, B. A. The human homologue of the *Aspergillus* nuclear migration gene *nudC* is preferentially expressed in dividing cells and ciliated epithelia. *Histochem Cell Biol***114**, 293-301 (2000).
- 10 Wang, Q. *et al.* RNA interference targeting CML66, a novel tumor antigen, inhibits proliferation, invasion and metastasis of HeLa cells. *Cancer Lett***269**, 127-138, doi:10.1016/j.canlet.2008.04.035 (2008).
- 11 Rao, W. *et al.* OVA66, a tumor associated protein, induces oncogenic transformation of NIH3T3 cells. *PLoS One***9**, e85705, doi:10.1371/journal.pone.0085705 (2014).
- 12 Rao, W. *et al.* OVA66 increases cell growth, invasion and survival via regulation of IGF-1R-MAPK signaling in human cancer cells. *Carcinogenesis***35**, 1573-1581, doi:10.1093/carcin/bgu070 (2014).
- 13 Wu, C. J. *et al.* Detection of a potent humoral response associated with immune-induced remission of chronic myelogenous leukemia. *J Clin Invest***106**, 705-714, doi:10.1172/JCI10196 (2000).
- 14 Taipale, M. *et al.* A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell***158**, 434-448, doi:10.1016/j.cell.2014.05.039 (2014).
- 15 Claerhout, S. *et al.* Abortive autophagy induces endoplasmic reticulum stress and cell death in cancer cells. *PLoS One***7**, e39400, doi:10.1371/journal.pone.0039400 (2012).
- 16 Umate, P., Tuteja, N. & Tuteja, R. Genome-wide comprehensive analysis of human helicases. *Commun Integr Biol***4**, 118-137, doi:10.4161/cib.4.1.13844 (2011).
- 17 Fouraux, M. A. *et al.* The human La (SS-B) autoantigen interacts with DDX15/hPrp43, a putative DEAH-box RNA helicase. *RNA***8**, 1428-1443 (2002).
- 18 Yoshimoto, R., Okawa, K., Yoshida, M., Ohno, M. & Kataoka, N. Identification of a novel component C2ORF3 in the lariat-intron complex: lack of C2ORF3 interferes with pre-mRNA splicing via intron turnover pathway. *Genes Cells***19**, 78-87, doi:10.1111/gtc.12114 (2014).
- 19 Modrek, B. & Lee, C. J. Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss. *Nat Genet***34**, 177-180, doi:10.1038/ng1159 (2003).
- 20 Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet***40**, 1413-1415, doi:10.1038/ng.259 (2008).
- 21 Graveley, B. R. Alternative splicing: increasing diversity in the proteomic world. *Trends Genet***17**, 100-107 (2001).
- 22 Blakeley, P., Siepen, J. A., Lawless, C. & Hubbard, S. J. Investigating protein isoforms via proteomics: a feasibility study. *Proteomics***10**, 1127-1140, doi:10.1002/pmic.200900445 (2010).
- 23 Torres, J. Z., Miller, J. J. & Jackson, P. K. High-throughput generation of tagged stable cell lines for proteomic analysis. *Proteomics***9**, 2888-2891, doi:10.1002/pmic.200800873 (2009).
- 24 O'Gorman, S., Fox, D. T. & Wahl, G. M. Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science***251**, 1351-1355 (1991).
- 25 Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol***26**, 1367-1372, doi:10.1038/nbt.1511 (2008).
- 26 Cox, J. *et al.* A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc***4**, 698-705, doi:10.1038/nprot.2009.36 (2009).
- 27 Yan, Y. *et al.* A novel mechanism of alternative promoter and splicing regulates the epitope generation of tumor antigen CML66-L. *J Immunol***172**, 651-660 (2004).

- 28 Gerhard, D. S. *et al.* The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Res***14**, 2121-2127, doi:10.1101/gr.2596504 (2004).
- 29 Trinkle-Mulcahy, L. *et al.* Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J Cell Biol***183**, 223-239, doi:10.1083/jcb.200805092 (2008).
- 30 Rothbauer, U. *et al.* A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol Cell Proteomics***7**, 282-289, doi:10.1074/mcp.M700342-MCP200 (2008).
- 31 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc***4**, 44-57, doi:10.1038/nprot.2008.211 (2009).
- 32 Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res***37**, 1-13, doi:10.1093/nar/gkn923 (2009).
- 33 Das, R., Zhou, Z. & Reed, R. Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. *Mol Cell***5**, 779-787 (2000).
- 34 Spector, D. L. & Lamond, A. I. Nuclear speckles. *Cold Spring Harb Perspect Biol***3**, doi:10.1101/cshperspect.a000646 (2011).
- 35 Nesic, D. & Kramer, A. Domains in human splicing factors SF3a60 and SF3a66 required for binding to SF3a120, assembly of the 17S U2 snRNP, and prespliceosome formation. *Mol Cell Biol***21**, 6406-6417 (2001).
- 36 Wen, X., Tannukit, S. & Paine, M. L. TFIP11 interacts with mDEAH9, an RNA helicase involved in spliceosome disassembly. *Int J Mol Sci***9**, 2105-2113, doi:10.3390/ijms9112105 (2008).
- 37 Gee, S. *et al.* Cloning of mDEAH9, a putative RNA helicase and mammalian homologue of *Saccharomyces cerevisiae* splicing factor Prp43. *Proc Natl Acad Sci U S A***94**, 11803-11807 (1997).
- 38 Arenas, J. E. & Abelson, J. N. Prp43: An RNA helicase-like factor involved in spliceosome disassembly. *Proc Natl Acad Sci U S A***94**, 11798-11802 (1997).
- 39 Martin, A., Schneider, S. & Schwer, B. Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. *J Biol Chem***277**, 17743-17750, doi:10.1074/jbc.M200762200 (2002).
- 40 Tsai, K. W., Tarn, W. Y. & Lin, W. C. Wobble splicing reveals the role of the branch point sequence-to-NAGNAG region in 3' tandem splice site selection. *Mol Cell Biol***27**, 5835-5848, doi:10.1128/MCB.00363-07 (2007).
- 41 Yoshimoto, R., Kataoka, N., Okawa, K. & Ohno, M. Isolation and characterization of post-splicing lariat-intron complexes. *Nucleic Acids Res***37**, 891-902, doi:10.1093/nar/gkn1002 (2009).

Figure 1

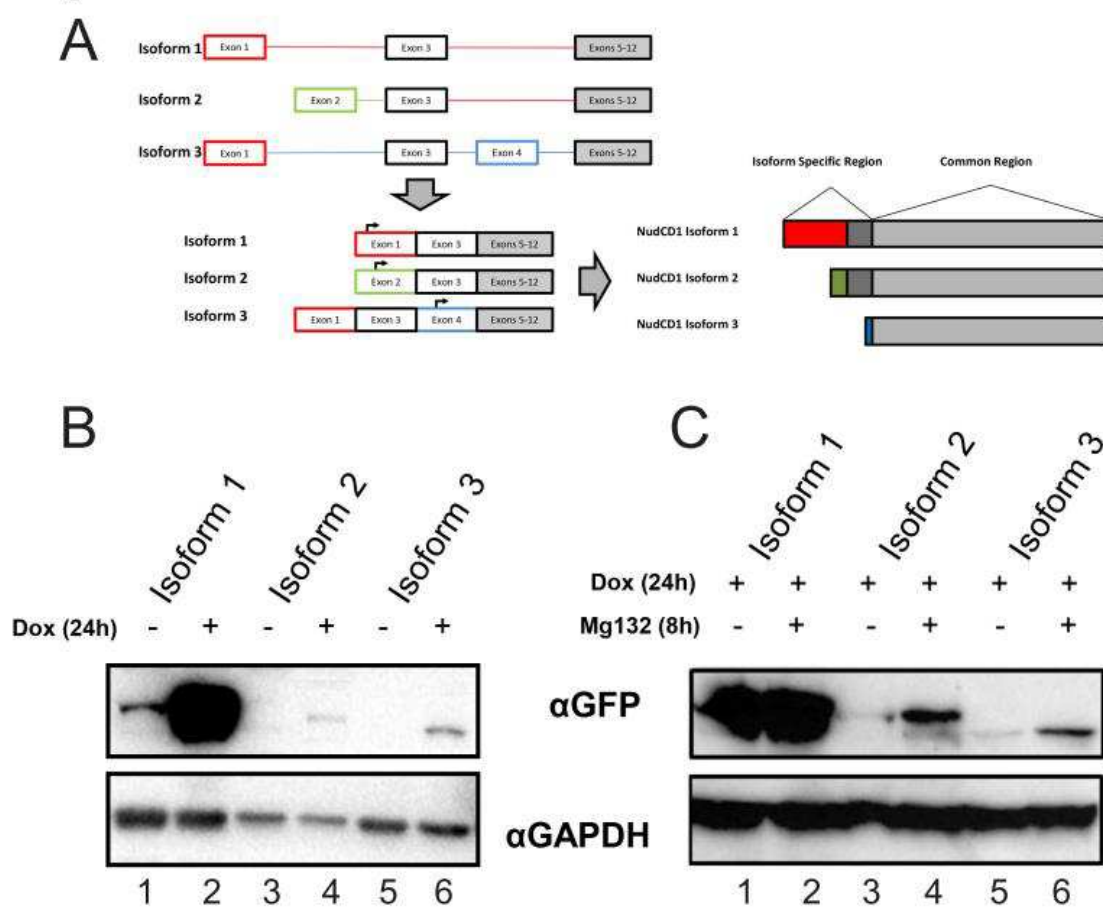


Figure 2

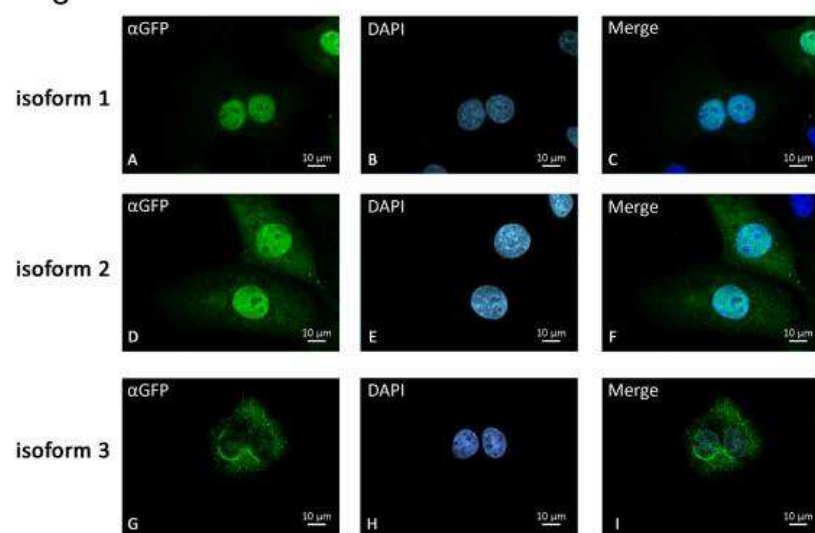


Figure 3

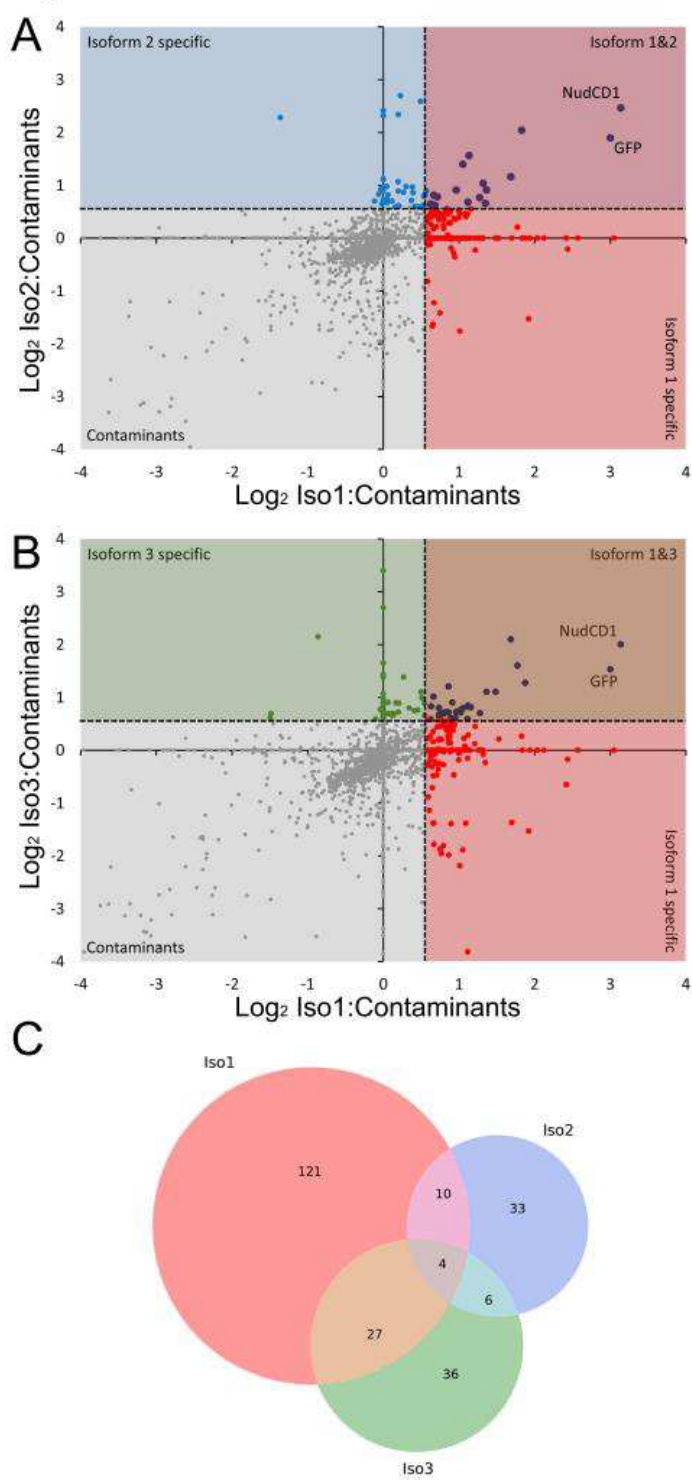


Figure 4

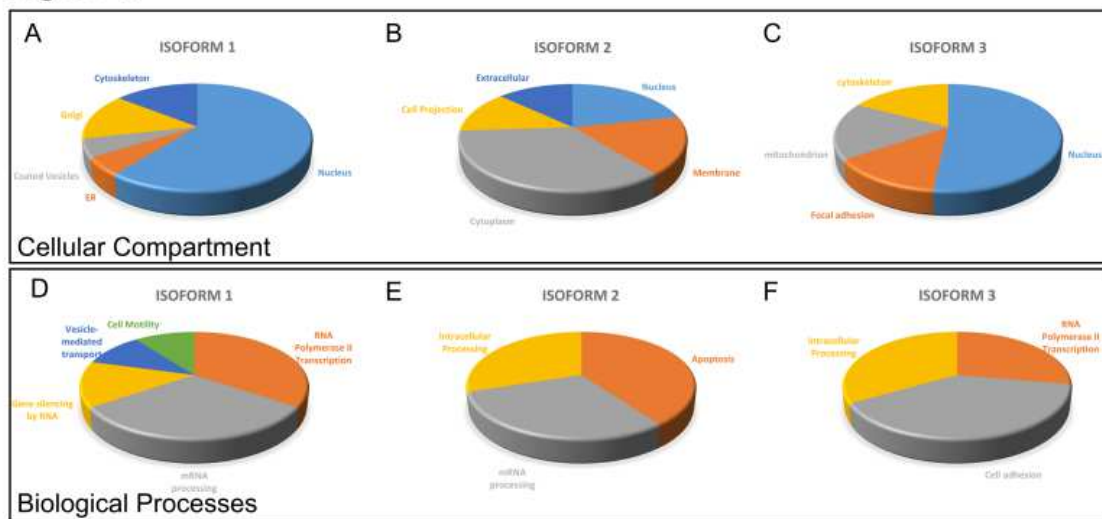




Figure 5

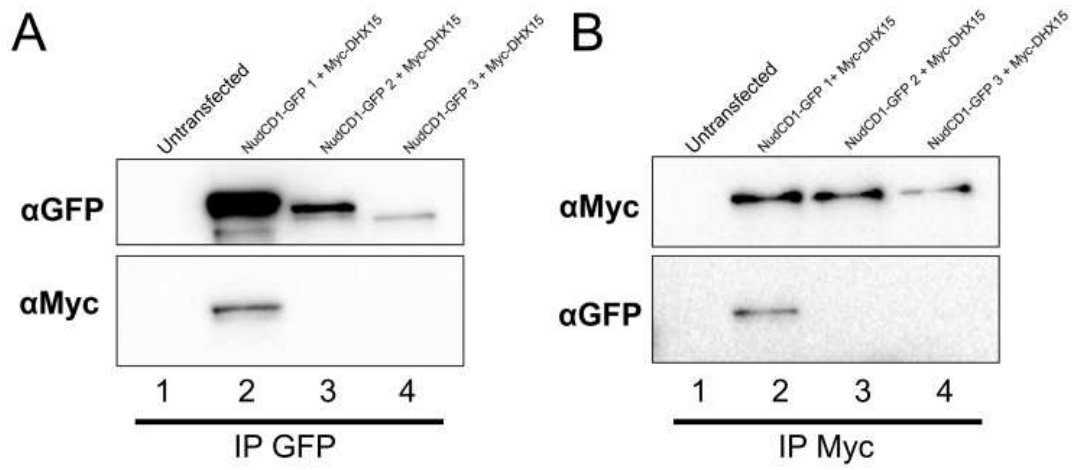


Figure 6

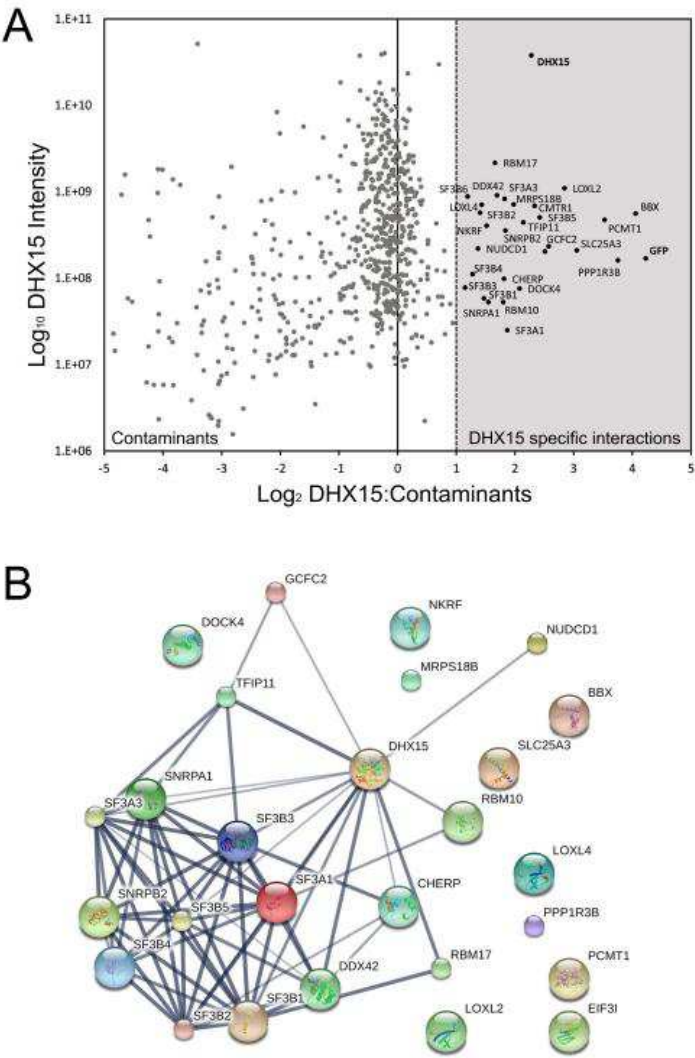
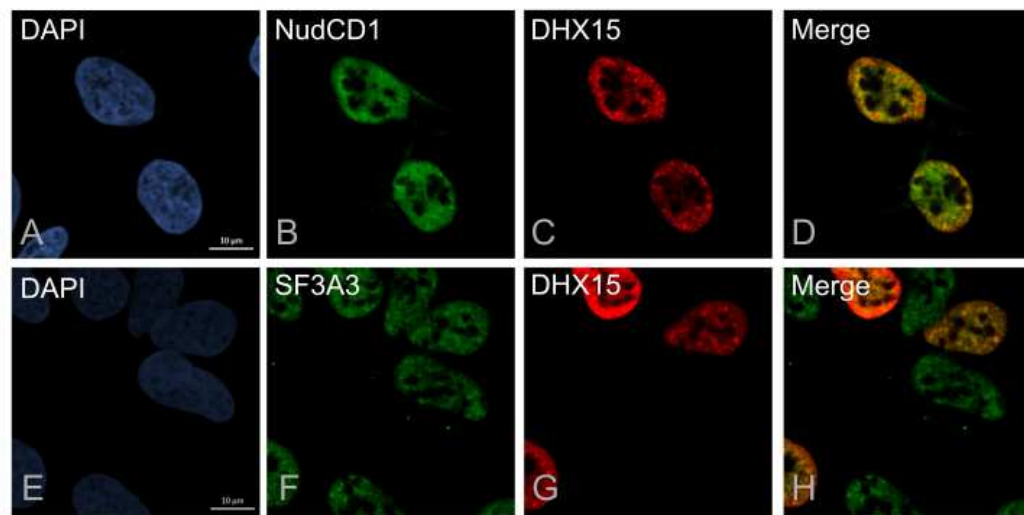


Figure 7



## 7.2 Amorces utilisées pour les RT-PCR

### NudCD1-1

*Forward* : 5'- CTAATTGCTCCCTACGGG -3'

*Reverse* : 5'- CAGCATTTAGCAGTGAGATAC -3'

### NudCD1-2

*Forward* : 5'- CTTTATCTCCAGGGTTGGAG -3'

*Reverse* : 5'- CAGCATTTAGCAGTGAGATAC -3'

### NudCD1-3

*Forward* : 5'- GTCATGTATACCCAGGACAC -3'

*Reverse* : 5'- CAGCATTTAGCAGTGAGATAC -3'

### GAPDH

*Forward* : 5'- TGATGACATCAAGAAGGTGGTGAA -3'

*Reverse* : 5'- TCCTTGGAGGCCATGTGGGCCAT -3'